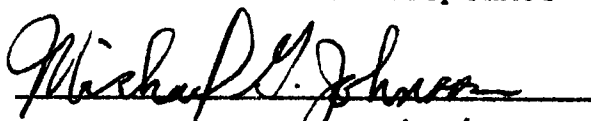
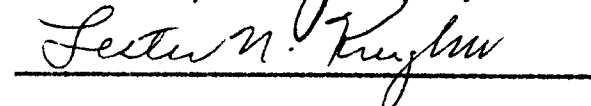
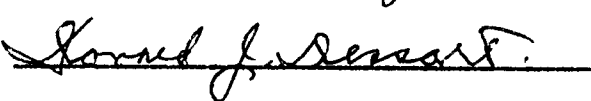


To the Graduate Council:

I am submitting herewith a dissertation written by Robert Rogers Wheeler entitled "An Investigation of the Relationship Between Category Development and Performance on a Standardized Test of Reading." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Education, with a major in Curriculum and Instruction.


J. Estill Alexander
Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:



Associate Vice Chancellor and
Dean of The Graduate School

AN INVESTIGATION OF THE RELATIONSHIP BETWEEN
CATEGORY DEVELOPMENT AND PERFORMANCE ON A
STANDARDIZED TEST OF READING

A Dissertation
Presented for the
Doctor of Education
Degree
The University of Tennessee, Knoxville

Robert Rogers Wheeler

May 1992

This dissertation is dedicated to those who have without qualification believed in me: my wife, Linda L. Wheeler; my mother, Nancy L. Vandermolen; my daughters, Sugar and Taffy. This dissertation is also dedicated to the memory of my father, Robert R. Wheeler, who pointed the way so very long ago.

ACKNOWLEDGMENTS

The guidance and instruction of Dr. J. Estil Alexander, who served as major professor, were central to the completion of this dissertation. His unerring advice and encouragement are sincerely appreciated. The author also wishes to thank his committee members, Dr. Lester Knight, Dr. Donald Dessart, and Dr. Michael Johnson for their advise.

Many others contributed directly and indirectly to the completion of this work and deserve more than the simple acknowledgment they receive herein: Mary Jane Hayes, Mary Kay Bordner, and Anna Lundy for their time and counsel; to Ann LaCava for her patient guidance; to Kathy Scott for the seemingly endless hours she worked in my place.

Gratitude is similarly expressed to the faculty and students of The Day School for their willingness to assist the author during the pilot study.

ABSTRACT

Background knowledge has long been considered to be central to reading. Categories of information may be one way of viewing the organization of background knowledge. Historically, categories have been considered to be fixed entities, developing only as additional similar exemplars are added. Recent research, however, challenged this. Five non-quantitative components or classifications of categories were identified in the literature. These were: (1) defining features, (2) prototypes, (3) natural kinds, (4) non-basic level categories, (5) taxonomies. It followed that reading development might be associated with category development of these.

Five questions, then, based on the relationship between each developmental component of categorization and both reading level and grade level were asked. These questions required measurement tools for reading development and categorization development.

The Gates-MacGinitie Reading Tests (1989) Level 2, Form K was used to measure reading development. The test was used to identify low readers (stanines one, two, and three) and high readers (stanines seven, eight, and nine) in first and third grade children in two rural east Tennessee elementary schools.

The literature supported the use of sorting tasks as measures of categorization. Therefore, five sorting tasks

were used to evaluate the categorizations of first and third grade low readers and high readers.

The following results indicated a pattern of development:

1. Moderate to strong associations between reading stanine levels and many of the categorization tasks for both grade levels.
2. A pattern of differences between the means of low readers and high readers.
3. A comparison of first grade low readers and third grade high readers which suggested that the effects of both reading level and grade level increased the magnitude of the group differences.

The results suggested that category development could be associated with reading development. This has implications for reading evaluation and instruction. Further research is needed for verification.

CHAPTER	TABLE OF CONTENTS	PAGE
I.	BACKGROUND AND ORGANIZATION OF THE STUDY	1
	Introduction	1
	Categories	2
	Need for the Investigation	11
	Research Questions and Rationale	13
	Methodology and Measurement	15
	Organization of the Study	15
II.	SURVEY OF THE LITERATURE	17
	Categories and Reading - An Introduction	17
	Challenges to the Classical Model	21
	An Alternative Explanation of Category	46
	Reading and Categories	49
	Summary of Major Findings	52
III.	RESEARCH PROCEDURES AND MATERIALS	55
	Introduction	55
	The Sample	56
	Data Gathering Procedure	58
	Sorting Task Development	63
	The Sorting Tasks	65
	Procedures for Analysis of Data	71
IV.	PRESENTATION OF DATA	75
	Introduction	75
	Overview of the Data	77
	The Questions	87
	Summary	98
V.	SUMMARY OF RESEARCH	106
	Introduction	106
	Conclusions	108
	Discussion	110
	Recommendations for Futher Research	115
	BIBLIOGRAPHY	117
	APPENDICES	126
A.	Letters of Consent	127
B.	Categorization Tasks	134
C.	Assumptions, Limitations, Delimitations	142
D.	Raw Data	145
	VITA	148

LIST OF TABLES

TABLE		PAGE
I.	Comparison of First Grade Low Reading Group (FLR) and First Grade High Reading Group (FHR)	77
II.	Correlation Between Stanine Score and Categorization Task Performance - First Grade Low Reading Group (FLR) and First Grade High Reading Group (FHR)	78
III.	Comparison of Third Grade Low Reading Group (FLR) and Third Grade High Reading Group (THR)	80
IV.	Correlation Between Stanine Score and Categorization Task Performance - Third Grade Low Reading Group (TLR) and Third Grade High Reading Group (THR)	81
V.	Comparison of All First Grade Subjects (FGS) and All Third Grade Subjects (TGS). . .	83
VI.	Correlation Between Grade Level Groups and Categorization Task Performance	84
VII.	First Grade Low Readers (FLR) and Third Grade High Readers (THR) By Categorization Task Performance	85
VIII.	Comparisons of Differences Between Means For First Grade and Third Grade, For Low Readers and High Readers, For Low First Grade Readers and High Third Grade Readers	86
IX.	Raw Data - First Grade Low Reading Group By Subject	146
X.	Raw Data - First Grade High Reading Group By Subject	146
XI.	Raw Data - Third Grade Low Reading Group By Subject	147
XII.	Raw Data - Third Grade High Reading Group By Subject	147

CHAPTER I

BACKGROUND AND ORGANIZATION OF THE STUDY

Introduction

Background knowledge has long been considered to be both complex and central to reading, involving both the subject matter under consideration as well as the reading act itself (Alexander, 1988; Cooper, 1986; Downing & Leong, 1982; Goodman, 1990; Kletzien, 1991; McKeown, Beck, Sinatra, & Loxterman, 1992; Nagy, 1988; Pflaum, 1986; Roller, 1990). Categories, and the words selected to represent them, may be one way of viewing the organization of background knowledge. The study of categories, then, may prove useful to the study of reading. Traditionally, the category has been treated as if it were a container (Lakoff, 1987). Category development simply involved the addition of instances of the category (exemplars) over time. Recent literature indicated, however, that this was not so. Category development, rather, appears to involve both qualitative and quantitative changes. At least five developmental changes, not predicted by the traditional model, have been identified. These were changes in the ability to use categorical information about definitive features, prototypes, natural kinds, non-basic level categories, and taxonomies.

Reading depends upon a transaction between the reader

and the text (Goodman, 1985). For this to occur categories both in the text and in the reader must match developmentally. It is critical, therefore, to examine the relationship between specific kinds of category development and the reading act.

Hence, the present inquiry examined the relationship between reading performance, as measured by the Gates-MacGinitie Reading Tests (1989), Level 2, Form K, and category development, as measured by five sorting tasks. Each sorting task was designed to reflect development in one of the five areas cited above.

Chapter I includes the stated introduction, a brief discussion about the traditional model of categories and the misconceptions that follow from such a model, research questions and the rationale for them, a brief statement of methodology, the purpose of the investigation, and the organization of the study.

Categories

Categories may be thought of as exemplars grouped on some theoretical basis resulting in perceived structural and/or functional similarity. As the theoretical foundation for the category changes, so, too, does the category (Medin & Wattenmaker, 1987). This appears to occur over time in an organized and predictable way.

Lakoff (1987) argued persuasively that this dynamic

conception of category appears not to have been the traditional one. Rather, categories have been treated (at least in schools) metaphorically, as if they were containers of fixed dimensions. Lakoff's argument is supported both by casual observation and by empirical research. Two illustrations are offered. First, at the casual level, it can be observed that information is often presented in isolation. Geography, social studies, anthropology, sociology, and history have long been taught as isolated disciplines (Benjamin, 1939; Taylor, 1989), as if the information for each was held in separate containers. Second, at the empirical level, reading researchers have observed that teachers, for the most part, ask questions calling for literal comprehension (Harris & Sipay, 1985), which may be thought of as exemplar recall. Responses to these "fact recall" questions are usually "correct" or "incorrect," suggesting the classical model of category-as-container with fixed and clearly defined boundaries. These literal questions suggest that categories are fixed dimension containers with varying numbers of equally good exemplars.

In a sense these two illustrations depict an entire model of learning that focuses on isolated containers of exemplars. . . as if the development of these categories involves little more than adding to the numbers of exemplars present in a given category.

The category-as-container metaphor supports at least five misconceptions regarding the nature of categories. These misconceptions have impacted the way background information has been presented for learning and cued for retrieval (Carnine, 1990). It follows that they may have affected the entire structure of reading programs, from reading material to testing material. These five misconceptions have formed the basis of challenges to the traditional model found in the literature. Hence they will be introduced here, followed by five research questions that were generated by them. Supporting theory and research will be offered in Chapter II.

1. The category-as-container metaphor supports the observation that the category is always determined by fixed boundaries (in other words, defining features). Defining features may be considered to be conditions necessary for the inclusion of an exemplar in a given category. These typically consist of some common property expressed as a predicate superordinate level noun (and attendant adjectives) preceded by a form of the copula as in "a dog is a four legged animal."

In contrast any feature of an exemplar that is often present in, but not necessary to, many exemplars of a category may be thought of as a characteristic feature. A characteristic feature is often expressed as an adjective as in "water is cool." "Cool" is not a necessary condition for

a liquid to be defined as water.

At first glance defining features appear to be rather easily understood, for many categories seem to be constrained by defining features. However, upon closer inspection, this appears not to be so simple a matter. For example, the word "bride" may be thought to represent a category of unmarried females. The condition for category membership is fixed by this boundary. Any fixed definitional boundary, however, presents certain problems, for this category as defined would include two year old girls. The boundary could be redefined to excluded young girls, but invariably, other problems arise. For example, another definition of "bride" might simply be "unmarried adult female." However, when a sufficiently general boundary such as this is established, the difference between this category and other categories may be trivialized. A "spinster" is also, after all, an "unmarried adult female."

Another example of a category that cannot always be easily constrained by definition is that of "dog." The child who has learned that the word "dog" represents an animal if and only if that animal has fur, has four legs, and barks, could never apply the word to an instance of a hairless silent dog who had lost a leg in an accident.

Hence, it must be concluded that the boundaries of the category are probably not easily understood by definition alone. Since definitive boundaries involve life experiences

and complex knowledge of superordinate levels of abstraction, definitive features are probably more useful to the older child. Consider the difficulty of trying to teach a very young child that many new words are represented by exemplars and that each of these has conditions for set membership. To teach by definition the concept of "table" the young child would need to have been taught the defining features of surface, elevated, leg, wood, etc. This would involve a dictionary style of learning. It does not appear that young children learn this way. However, over time children come to understand and use defining features. The ability to use defining features, then, would appear to be developmental.

2. The category-as-container metaphor supports the observation that all exemplars within a category ought to be equally good. If defining features determined category membership, each exemplar in the category would be essentially the same. Each exemplar (and the word that symbolizes it) would be equally likely to be retrieved and would be equally representative of the category.

If this were the case, young children would find the word "canine" as useful as the word "dog." They would find an ostrich and a robin equally good examples of the category "bird." Hence, it seems that each exemplar is not equally likely to be retrieved nor equally representative.

3. The category-as-container metaphor supports the

observation that category development is quantitative only. Accordingly, the category would develop only in terms of the number of similar entities and representative words.

Development, however, may be thought of as any quantitative and/or qualitative change in the structure or function of a category or in the relationship between categories that is correlated with age. According to the classical view, no qualitative development can be anticipated.

The distinction between categories of natural kinds and categories of artifacts has proved a useful way of illustrating that, in fact, the quality of categorical information does develop over time.

A natural kind category, for the purposes of this discussion, is a category of things the origins of which are unknown. A natural kind is often expressed as a noun, as in "cat." Natural kinds may be contrasted with categories of artifacts. An artifact may be thought of as anything the origins of which are known, as, for example, in the case of "chair." Changes in the properties of artifacts produce category redesignations. If walls are removed from a house it becomes a patio. If the back of a chair is removed, it becomes a stool. If the head of a mop is removed, it becomes a handle.

Changes in the properties of natural kinds, however, do not result in category redesignations. A person without

legs is still a person. A dog without a tail is still a dog. A tree without leaves is still a tree.

It appears, then, that there are qualitative differences between categories not predicted by the category-as-container metaphor. It appears, also, that the ability to utilize these differences increases with age. To a small child, after all, the origins of almost everything are unknown.

4. The category-as-container metaphor supports the observation that categories are of a single level of abstraction. Children should not, therefore, develop superordinate level, basic level, and subordinate level categories.

However, these levels have been identified in the literature. The basic level is treated as a vertical category designation based on inclusiveness and assigned to any category of things that arouses the senses and/or names things more inclusive than subordinate level and less inclusive than superordinate level. It is often expressed as a single instance of a conceptual collective noun (for example, forest) and generally represented by a single word, as in "tree." The subordinate level is considered to be a vertical level designation based on inclusiveness and assigned to any category of things that are specific in nature, that name specific kinds of things, and that are less inclusive than basic level things. The subordinate

level is often expressed as nouns and adjectives in a conceptual unity, as in "bicycle-built-for-two" or in "pine tree," or as a particular kind of a basic level exemplar, as in "Cadillac." The superordinate level may be thought of as a vertical level designation based on inclusiveness and assigned to any category of things that are collections and are more inclusive than basic level. It is often expressed as a collective noun, as in "forest." These differences appear to exist between categories at varying levels of abstraction. These are not quantitative differences. Rather, they are qualitative differences that cannot be accounted for by the category-as-container metaphor. Moreover, they appear to be developmental.

Anyone familiar with a toddler is likely to have observed that the word "tree" is more often uttered before "forest" or "sugar maple." Similarly, most children use the word "car" before "transport" or "Mercedes."

5. The category-as-container metaphor supports the observation that categories will be described at all stages of development in terms of taxonomies, or categorizations based upon enumerations of features of perceived structural or functional similarities. These features or attributes may be thought of as properties of category exemplars. They are often expressed as adjectives, as in "the tall brick building."

This view suggests a rule regarding category formation

that should be universal. However, it appears not to be. Categories often suggest to children relationships rather than properties. Upon cue, observation suggests that children are as likely to talk about relationships as they are to list features. For example, there is nothing about a dog and a toy bone that is alike, yet children will group them together, evidently relying on the nature of the relationship rather than features of the exemplars themselves.

Thematic groupings, then, are groupings of categories based on relationships; often expressed through the use of transitive verbs and direct objects, as in, "The dog chased the cat." It seems likely, then, that there may be an inverse relationship between the use of themes and the use of taxonomies in categorizations.

It follows that some categorizations are not represented taxonomically in all instances and at all times. Rather, at least for young children, they may often be represented thematically.

Each of these five misconceptions followed from the popular use of the category in which it was considered to be a container. The container was popularly viewed as an entity of fixed proportions - empty until filled. In each case, however, this conception of category does not appear to be universally correct.

The five areas of misconception concerned defining

features, prototypes, natural kinds, non-basic levels, and taxonomies. Each of these areas appears to develop in qualitative as well as quantitative ways. Each of these appears to affect the structure and formation of background knowledge necessary for reading and, as a result, the selection of exemplars and representative words. No research on the subject of category development, however, has related it to reading. Hence, these five areas will serve as the basis for the formulation of the research questions and, later, as the basis for a survey of the literature for this study.

Need for the Investigation

It has been suggested that the classical interpretation of the development of category involves the use of the container metaphor. It has been further suggested that this may not be accurate. A dynamic developmental approach to categories has been suggested. Since this may be the structure of the background knowledge which is critical to reading, it is important that research address the relationship between category development and reading development.

The issue of reading development and text development has been confronted by several authors (for example, Alexander, 1983; Alexander & Heathington, 1988; Dolch, 1939; Harris & Sipay, 1985; Pflaum, 1986). Some have addressed

specific cognitive structural and functional changes instrumental in the development of the ability to read (Alexander, 1983). A very few have briefly touched on the subject of categories and reading (Gillet & Temple, 1990; Carnine, 1990). Neither in the literature on category nor in the literature on reading, however, has an investigation directly addressed the intimate relationship between category development and reading.

Perhaps this should not be surprising. Goodman (1985) noted that the interaction among reading and related disciplines has been limited. Most of the theory and research in categories has been done in related areas and, hence, may have been overlooked by reading and other disciplines (Keil & Kelly, 1987; Medin & Barsalou, 1987).

Misunderstanding the development of categories potentially corrupts both the creation of age-appropriate texts and the evaluation of reading disorders. This potential for problems creates a need for extensive research in the area. Research in category development and reading promises much. Understanding this relationship could lead to a better overall understanding of reading development, guide in the process of the selection of words for texts, and assist in the evaluation of reading progress.

An examination of the relationship between reading and categories is clearly needed. One purpose of this study, then, is to describe the relationship between categorization

development, as measured by five categorization tasks, and reading performance, as measured by one standardized reading test, the Gates-MacGinitie Reading Tests (1989), Level 2, Form K. Another purpose of this study is to describe the relationship between categorization development, as measured by five categorization tasks, and grade level. These five categorization tasks were generated by the five research questions.

Research Questions and Rationale

Categories appear to develop in ways not predicted by the classical model. This is supported by the review of the research which will be presented in Chapter II. Hence, for the purposes of this discussion, a dynamic developmental view of category has been adopted.

Categories and the words that represent them serve as structure for background knowledge (Barsalou, 1986, 1987; Dale, 1972; Keil & Kelly, 1987; Markman, 1989; Matlin, 1989), and as such appear to be essential to reading. The development of reading would seem to coincide with the development of categories. This relationship has not, however, been demonstrated. In general, then, the research questions in this study focus on describing the relationship between developmental changes in categories and reading performance. Specifically, each of the questions inquire about a possible relationship between age-related sorting

task (categorization) and reading test performances.

The rationale for all of the questions is the same. In each case, the development of the aspect of categorization under consideration may be viewed as a constraint upon successful reading performance. These questions, then, focus on the relationship between the development of categories and reading performance. If a relationship is indicated, it may have implications for both reading programs and reading assessment. The five research questions generated by this discussion are as follows:

- I. What is the relationship between the identification of defining features and reading test performance or grade level?
- II. What is the relationship between the production of prototypes and reading test performance or grade level?
- III. What is the relationship between the identification of natural kinds and reading test performance or grade level?
- IV. What is the relationship between the identification of superordinate level categories and reading test performance or grade level?
- V. What is the relationship between the production of taxonomic descriptions of category exemplars and reading test performance or grade level?

Methodology and Measurement

The sample of 33 (16 first grade children and 17 third grade children) was selected by means of the Gates-MacGinitie Reading Tests (1989), Level 2, Form K from a population of 89 rural East Tennessee school children. The subjects within the sample were given five sorting tasks thought to indicate development in each of the identified areas and included only those children who scored in the bottom three and top three stanines in each grade level. It was then possible by means of one-tailed t -tests, Pearson Product Moment coefficients of correlation, and probability tests to compare the categorizations of the better readers with the poorer readers within each grade level, between grade levels, and the high third graders with the low first graders.

Organization of the Study

The present study is divided into five chapters. Chapter I includes introductory comments, a discussion of categories, the research questions and the rationale for those questions, a brief statement of methodology, and the purpose of the investigation.

Chapter II consists of a discussion of literature related to the subject of categories and reading.

In Chapter III an explanation of the research procedures is given, and a rationale for these procedures is

offered.

The results of the research are submitted and explained in Chapter IV.

Chapter V contains a summary of the investigation, conclusions, a discussion, and recommendations.

The Appendices consist of permission letter styles, sorting tasks, limitations and delimitations, and assumptions.

CHAPTER II

SURVEY OF THE LITERATURE

Introduction

A survey of the literature indicated that categories have often been treated as fixed entities, changing only in quantity. Recent literature argues that this is not the case. Categories also develop in qualitative ways. Further, the literature implies that this development may be related to reading development.

Lakoff (1987) maintains that the classical view of category owes its existence to an objectivist view of reality. This position holds that all things have a meaningful existence apart from human experience. Things, then, are classified according to the properties of the things themselves rather than according to the properties of the organism doing the classifying. The categories people use are the result of mapping this external world onto the internal mental world. All elements must transfer if robust meaning is to be the result. From this point of view things are seen to be naturally categorized according to some principle of maximum discontinuity. Necessary and sufficient conditions determine points of discontinuity between categories (Kiel, 1987) and are derived from the things being categorized. All exemplars ought to be at all

times during the life of an individual (1) represented by definitive features, (2) equally good (without prototype), (3) functionally equivalent from one category to another, (4) invariant with regard to abstraction level, and (5) cues for lists of features and for taxonomies. Only quantitative changes might be anticipated. The literature indicates that this is not the case.

According to the classical model, then, groupings of things are real; separate from human experience. Once conjoined to mind, categorization has occurred. For many years this position was not questioned. After all, this was consistent with the history of Western thought (Lakoff, 1987), consistent with Western theology (Ozman & Craver, 1986), more or less measurable, and seemingly consistent with subjective human experience. Each category was defined in terms of its' limits, and its' limits were defined in terms of the category. There was an elegant simplicity to this, and it would have been surprising if reading education had not adopted this point of view. There is, however, reason to believe that this tautology (some of which was explored by Chomsky, 1959) like all others, was both misleading and inadequate.

If things in the world are real, then the categories to which they seem to adhere are real, and there ought to be a one to one relationship between categories in the world and categories in the mind (Lakoff, 1987). From this it should

follow that each individual holds in mind exactly the same categories and that these categories in mind will enjoy similar structural and functional characteristics. Information retrieval should, then, be uniform over time. Experience (if not empiricism) reveals that this is not the case.

For all of its' shortcomings, the classical model still appears to impact Western thought. Lakoff (1987) asserted that this point of view is so thoroughly shuffled into the philosophical deck of western man that it will not be easily discarded. Historical antecedent and ubiquitousness, however, insufficiently account for the popularity of the objectivist argument. For a particular view of the world to be so enduring, it must enjoy some measure of success in describing the world.

There are indications that this point of view enjoyed such success in education. Objects and actions from the world seemed to be thought of as placed in the child's mind by means of print. One spoke of a child "receiving" an education (Skinner, 1953), as if external physical objects were moved into some mental storage unit, and the implication was that the storage unit was something similar to the category.

Lakoff (1987) observed that the objectivist argument remains even today a powerful force in education. Identification of plants and animals, for example, is an

attempt to recognize and categorize them based upon objective attributes of the plants and animals themselves. Many categorizations based upon the classical conception, however, simply do not work. Lakoff (1987) noted, for example, a particular animal that looked like a zebra but did not share the same internal structure. The classical model could not be used to classify this, for it provided no mechanism to account for the weighting of disparate features. If things cannot be classified, then they must initiate new categories. But for each thing in the world to potentially receive a new and separate classification, each thing must be unrelated. Even the classical model, however, predicts categories with multiple exemplars.

Reading was taught as if most of what needed to be known was contained, somehow, in the print itself. Many of the data driven models suggest this. Gough (1985a), for example, held that letters were assembled into meaningful units. This suggests that the letters somehow contained the meaning. However, even Gough (1985b) himself later admitted that his early model of reading had been incorrect. Arguments put forth (Chall, 1967) in support of phonics instruction may be seen as favoring this position.

This point of view characterized mind as somewhat passive. Mead (1962), however, held that mind was actively responsive to objects. Gibson (1966) also discussed our active response to objects. Mind, then, for Mead and Gibson

existed in an active relationship to things. This portrayal of mind as an organizing force is supported by current views of category and challenges the classical position.

The result of these early and general challenges is seen in the literature. A developmental view of categories useful to the production of reading curriculum and reading assessment and to the production of reading theory and reading research has emerged. It is appropriate, therefore, that these challenges to the classical position be presented. These challenges form the basis of a new model of the structure of background knowledge necessary for meaningful reading. Therefore, the classical model will be confronted from the position of similarity judgments, category stability, category development.

Challenges to the Classical Model

Challenges to the traditional model of category have often been based upon similarity and stability. It has been shown that traditionally things were considered to be defined by sets of inherent properties, and categories came to be viewed as groupings of things that shared these properties. According to this point of view, things are included in categories by means of judgments of similarity. Also, according to classical theory, categories should be stable. From moment to moment the information retrieved from a given category should be constant.

The classical position further suggested that only quantitative changes occur in categories, and that categories develop only quantitatively. While the issues of similarity and stability are important in a general sense to an understanding of reading, the issue of development is critical.

The five challenges cited above are supported by the literature generated by these issues and by the theoretical work of Piaget (Waller, 1977), for these developmental changes appear to occur during the transition from later preoperational thought to concrete operational thought.

Issue One: Similarity

The classical position treats the category as a container of similar things. Challenges to the classical model, then, are challenges to a similarity based model. The classical position explains neither similarity itself nor why this similarity is often undefined. If similarity is undefined, it is difficult to see on what basis similarity determinations would be made. After all, at some level any two things could be judged similar (Medin & Wattenmaker, 1987). An orange and Lake Michigan would then be the same since both contain fresh water, have irregular surfaces, contain living tissue, occupy space, etc.

A gross weakness in similarity models and their attendant theories, then, is that they cannot account for

the very similarity upon which they seem to depend. When an effort is made to do so, similarity theories become circular (Medin & Wattenmaker, 1987): things are placed in a category because they appear to be similar and are similar because they are placed in the same category.

Another weakness is that similarity alone does not account for the development of defining features, prototypes, qualitative differences between categories, levels of category abstraction, or why young children seem to prefer to categorize by theme (Medin & Wattenmaker, 1987; Kiel, 1987) rather than by taxonomy.

Issue Two: Category Stability

The classical model predicts qualitative stability in both storage and retrieval of information. Change is predicted only in the quantity of exemplars captured within a category. Accordingly, definitive features, prototypes, all category exemplars, levels of abstraction, and taxonomies should remain constant. However, recent literature suggests category instability. This instability may take several forms. Included are the following:

1. Property Interaction. Category inclusion may be determined by the interaction of variables. A fully dressed man who jumps into a swimming pool might be categorized as a clothed male swimmer (Medin & Wattenmaker, 1987). However, the

interaction of the variables suggests that the man is unstable, drunk, or in the process of saving someone from drowning. Property interaction defines the category here.

2. Context Dependent Information. The context of information also appears to influence the information retrieved from categories (Feldhaus & Honeck, 1989; Harnad, 1987; Medin, 1986). It was found that context determined why, for example, paper currency is not commonly cited as a prototypical exemplar for the category of things that are flammable (Barsalou, 1987). Barsalou (1987) examined other context effects. It was observed that when people read the word "frog," "eaten by humans" remained inactive. However, when primed by printed material describing frogs as served in French restaurants, "eaten by humans" becomes active.

Barsalou (1985), then, treated many categories as graded structures (any categorical structure which includes a continuum from best example to worst example). He showed that ideals, central tendency, and frequency of instantiation were possible determinates of these graded structures, which most certainly are neither classically structured nor classically stable.

Moreover, these structures themselves do not react uniformly, for even the impact of these context dependent variables depended upon whether the category was a common taxonomic one or a goal derived one.

Another source of contextual variation is reminders. Ross, Perkins, and Tenpenny (1990) provided subjects with three printed booklets. The first contained a person's name and a fact about that person. The second booklet placed that person in a group and gave a fact about that group. The third booklet asked that new instances be categorized. The results of the study indicated that reminders significantly affected categorizations ($p < .01$).

3. Context Independent Information. Another source of variation in retrieved category information is context independent information - that information that has been incorporated in a concept often enough to achieve a contextually independent status (Barsalou, 1987). Barsalou (1982) found that "unpleasant smell" was activated automatically on all occasions without regard to context when subjects were exposed to "skunk."
4. Recency of Experience. Still another source of variation in categorical information retrieval was

accounted for by the recency of the experience (Barsalou, 1987). Barsalou (1987) examined protocol overlap over a period of a few weeks (context was held constant). Two experiments produced an average overlap of about .55. Only a bit more than half of a subject's protocol regarding a category in one session overlapped with her/his protocol in another session. In other words, an individual shows substantial change in the information retrieved over time.

5. Goals. Barsalou's work (1985, 1986) suggested that categories of information are often assembled in a tentative fashion dependent upon the goals of the individual and the kinds of information available. If, for example, a category of clothing one might wear to impress a supervisor was formed, it might well not include the same items as a category of clothing one might wear to impress a date. Barsalou's (1987) assessment of the instability of graded structures over time concluded in the contention that whatever information is in fact retrieved from a category in the presence of given cues (which is most probably only a fractional part of the total information available) is goal dependent.

Issue Three: Category Development

While category instability predicts transient events without regard to age, developmental shifts are fairly enduring, age related patterns of categorization. Five of these enduring changes are not predicted by the classical model. These are of particular interest as they formed the basis for the research questions already given. They include the abilities to distinguish between and/or use: (1) defining features and characteristic features (2) prototypes, (3) natural kinds and artifacts, (4) superordinate, subordinate, and basic level categories, (5) thematic relationships and taxonomic relationships.

Defining Features

The classical model of category suggests that the boundaries to word meanings - as they represent categories - are fixed. These fixed boundaries are determined by defining features. Rosch (1978) observed that these criteria are, in fact, present in virtually all formal definitions.

Defining features often call for an understanding of function, which children often lack. A wagon, for example, is a wheeled vehicle used to (or functions to) move objects. If the conditions of being a wheeled vehicle and moving objects could somehow be demonstrated to be defining, it is unlikely that young children could use this information to

categorize the word "wagon," as they would have no basis for doing so. They would have no category describing the function of wheel or vehicle or any notion of function itself, for that matter. While it is arguable that in fact young children do have trouble with categorizations (as observed in both the overextension and underextension of the meanings of certain words), it is not arguable that somehow they do eventually come to categorize "wagon" and many other things. Although the conception of defining features may be useful in a formal environment to determine whether or not an exemplar belongs to a given category, it does not appear, then, to be solely responsible for categorizations (Dale, 1972; Langacker, 1987). The adult conception of defining features does not appear to be central to young children's considerations.

Kiel (1987) found that there was a consistent domain specific shift away from the use of characteristic features as children grew older. Children were asked whether a man who was kind and brought them presents and was a friend of the children's parents was an uncle. This was considered to be a high-characteristic/low-defining description. Another description involved a man who was the two-year-old brother of the children's mother. This was considered to be a high-defining/low characteristic description. Keil found that young children did not seem to regard any single feature as essential, or defining, for the categorization of exemplars

of "uncle". Similarly, given a peninsula, young children were likely to classify it as island if it had three of the following four characteristic features: people in grass skirts, palm trees, buried treasure, and/or a sandy beach. Keil (1987) drew further support from anecdotal evidence. A young child, for example, may only insist that one must leave home to have a vacation.

Markman (1989), noted that Tversky (1985) found young children almost always based inferences on similar looking rather than similar functioning objects. Tversky's work suggested that four year old children will typically group fire engines with apples because they are both red, rather than fire engines with cars because they are both vehicles. Their evaluations, then, were based on characteristic rather than defining features. Markman's (1989) work supports this.

Schank, Collins, and Hunter (1986) used a computer model to examine the problem of defining features. They found that it was impossible to tell in advance of contextual information, what features of an event were defining. Given a computer program that is directed toward the goal of identifying hijackings, they argued, and given that all hijackings to which it had been exposed had occurred in Cuba, it would be unable to deal with one that occurred in Libya. The new information would either cause the alteration of necessary and sufficient conditions - a

contradiction of the classical model - or it would not be considered a hijacking.

Dale (1972) identified a developmental shift toward increasing differentiation of categories (which might also be seen as a developmental shift away from characteristic features). In a longitudinal diary study Mervis (1987) found progressive changes in one child's use of the category "mixer." At fifteen months the following exemplars of the category "mixer" were acceptable to the child: an electric mixer, a hand pump and bucket of water, and a sewing machine. Mervis (1987) felt that these exemplars were different from adult exemplars because of both quantitative and qualitative changes in the child's evaluation of attributes. Changes, or shifts, like these are not predicted by the classical model.

It has been argued that a strict set of defining conditions cannot consistently account for categorizations. However, it has also been suggested that there is a developmental shift toward the use of these defining conditions. The classical position does not predict this developmental shift.

Prototypes

The classical conception of category predicts sets of highly similar exemplars. Words that represent them should be nearly synonymous. However, evidence offered above

suggested that category exemplars frequently exist in relationship to some set of ideal properties. Hence, there often may be a best example, or prototype (Matlin, 1989), of a category. This prototype is not necessarily "nearly synonymous" with other examples of the category. The category "bird" is not, then, equally well represented by robins, penguins, or ostriches.

This is not intended to suggest that one and only one exemplar differs from all others. Wittgenstein (1953) maintained that exemplars varied widely within a single category. Wittgenstein further held that family resemblance is one way to describe these category variations. Family resemblance can be thought of as any relationship among the exemplars of a category in which a distribution of attributes overlaps in such a way that all exemplars are linked attributionally to some other exemplars. No exemplar, however, is linked attributionally to all other exemplars.

Witgenstein examined the category game and noticed that board-games, card-games, Olympic games, etc. all had something in common with other instances of games, but no single element was common to all instances of the category game. Family resemblance, then, suggests that each exemplar shares one or more attributes with one or more of the other exemplars.

Rosch (1973) saw this organization as graded from best

example, or prototype, to worst example. Rosch (1973) selected eight colors judged cross-culturally to be prototypical (green, for example). There was, according to Rosch, a prototypical blue that best illustrates blue, that can be learned more easily and remembered longer. After learning to criterion, the Dani people of New Guinea who have no word labels for such colors, could have been expected to answer correctly by chance 1 time in 8. Instead, the overall ratio of correct answers was 7.2 to 8. A subsequent two-way analysis of variance for number of errors per stimulus indicated that the main effect for the focal dimension was significant ($p < .01$), suggesting the power of prototypicality.

Rosch's work with the Dani people established the legitimacy of the concept of prototypicality. In addition, Rosch (1978) found that prototypicality affected the following dependent variables:

1. Reaction Time. Rosch, Simpson, and Miller (1976) had subjects learn artificial categories with defined prototypes. Subjects were then asked to judge statements about category membership. Reaction time was found to be a function of prototypicality.
2. Speed of Learning. Rosch et al. (1976) found that prototypicality predicted the speed of learning of categories.

3. Order of Development in Children. Rosch (1973) found that reaction times for good and poor category exemplars for ten-year-old children were notably different than for those of adults. Rosch held that this suggested that prototypes were learned earlier than other category members.
4. Order of Retrieval. Rosch et al. (1976) demonstrated that the order of retrieval of exemplars from a category was from prototypical to non-prototypical.
5. Priming. Rosch et al. (1976) showed that the degree of prototypicality determined whether information provided about a category name before a matching test facilitates or inhibits responses.
6. The Use of Categories in Natural Language. Rosch (1978) cited evidence in language of the graded structure of categories. The use of the words "almost" and "virtually" indicate degree of representativeness. Another indication of representativeness is exemplar/category name substitutability. Some category members may be used interchangeably with the category name while others may not. For example, Rosch (1978) observed that the word "sparrow" may be substituted for the word "bird" in a sentence about birds sitting on telephone wires. The word

"turkey," however, cannot be substituted, though it clearly belongs to the category "bird."

Rosch's (1978) previously discussed work with prototypes illustrates the potency of the operations surrounding prototypicality. Rosch's assertions have been corroborated by many. Posner and Keele (1968) for example, found that prototypical dot patterns were correctly classified more often than control patterns. Mervis and Pani (1980) taught children categories using good and poor examples. Good examples (more prototypical) resulted in appropriate generalizations. Otherwise the categories formed were much narrower and the generalizations less appropriate.

Kelly, Bock, and Keil (1986) observed that the impact of prototypicality is so great that subjects would change the serial position of words such that prototypes were mentioned before non-prototypes ($p < .05$). In the same study subjects rated sentences more natural when prototypes were mentioned before non-prototypes in conjoining phrases.

Rosch's early work (1973) was seductive in its simplicity, however, and led some to feel that prototypicality was more than a description of category structure; that it was an explanation of category function. One consequence of the failure to isolate structure from function was a confusion in the use to which the concept of prototypicality was often put. Prototypicality could be

thought of as an average of features, or as an ideal, or as a goal, etc.

Medin and Wattenmaker (1987) and even Rosch (1987) herself in her later work implied that prototypicality may be more structurally descriptive than functionally explanatory. Medin and Barsalou (p. 463, 1987) discussed prototypes descriptively as containing:

. . . the characteristic attributes of a category's exemplars, namely, attributes probable across category members, but that are neither necessary or sufficient for category membership.

The process of category inclusion, therefore, has come to be seen as based upon something other than a similarity comparison of item to prototype (Geeraerts, 1989), or the most frequently retrieved exemplar, or an ideal, etc.. It may, in fact, be a composite that is as much a result of categorization as it is a cause of categorization.

Prototypicality, therefore, may not be elemental. Rather, it may consist of more than one component. Two predictors of prototypicality, or goodness of exemplar (out of others of less interest here) were isolated (Barsalou, 1985): frequency of instantiation and ideals. Barsalou (1985) found an ideals exemplar-goodness correlation of .70 and a frequency of instantiation exemplar-goodness correlation of .72 for goal derived categories. The correlation, however, between exemplar-goodness and central tendency was much lower ($r = .38$). The frequency of

instantiation of an exemplar or the ideal nature of an exemplar were found to be better predictors of prototypicality than was central tendency.

Although prototypes may, then, be seen as complex, they have been dealt with in the literature for the most part as a unity that develops over time and that, once instantiated, becomes a valuable cue for information retrieval. Rosch (1973) displayed sentences naming a category and an exemplar. Sentences such as "A robin is a bird" or "A pear is a fruit" were rated true or false. The reaction time for sentences in which the exemplar was central to the category rather than peripheral to the category were significantly shorter for child subjects - $t(19) = 2.63, p < .01$. They were also significant for adult subjects - $t(23) = 6.54, p < .001$. However, the reaction times for adults were less than half those for children - 1011.67 milliseconds for adults, 2426.45 milliseconds for children. This suggests both a pattern of reaction to prototypes and a pattern of development.

Prototypes, then, appear to become increasingly useful tools for information retrieval as children grow older. Underwood and Underwood (1987) studied children's organization and retrieval of information in a set of experiments. In the first experiment children (10 years of age) were less successful organizing information structures that resembled graded categories than adults (25% success

rate for children and 59% success rate for adults). In the second experiment, children were asked to identify prestructured information. Over half (52%) of the children failed to recognize hierarchies. Rather, they altered the information so that it formed a list rather than a graded structure. In a third experiment, children accessed information from various kinds of information structures which might be present in categories. A tabular form somewhat similar to a list was again the preferred structure. Children accessed information correctly 49% of the time from tables. Adults accessed information correctly 78% of the time. The Underwood and Underwood (1987) studies clearly show a developmental preference for prototypically organized graded structures.

The above research suggests, then, that all exemplars are not equally useful. This is not predicted by the classical model of categories.

Natural Kinds

Many categories of things can be thought of as either natural kind or artifact. Natural kinds are those things found in nature, the origins of which are not directly observable. They may be contrasted with artifacts for which the origins are known. Some distinguishing differences have been observed between the treatment of those things considered to be natural kinds and those things not

considered to be artifacts.

Keil (1987) inspected the nature of natural kinds and found that subjects resisted the reclassification of an exemplar placed in a natural kind category. Perceptual changes ordinarily do not affect categorization of natural kinds. Suppose a dog's larynx could be surgically replaced with a cat's larynx such that a meowing dog was produced. Then suppose that the dog's legs were replaced by the cat's along with a portion of its' central nervous system such that it could climb trees. Many people would still say that the animal in question was a dog; a strange, meowing, tree-climbing dog, but a dog nonetheless.

Perceptual differences in artifacts, however, do result in the movement of exemplars from one category to another. If subjects are asked what a chair with the back removed and the legs extended has become, they are likely to claim that it is no longer a chair. Rather, they are likely to claim that it is now a stool. Although part of this distinction may be based on function (Barton & Komatsu, 1989), it is the matter of the distinction itself that is important here.

Artifacts, Markman (1989) observed, are not usually expected to share many features besides the obvious ones, and, hence, do not support as many inferences. Artifacts are bound perceptually to their categories, and their origins are directly observable or known. Ladders and houses have known origins, for example, but a ladder without

rungs is not a ladder, and a house without walls or roof might be a deck, but it is certainly no longer a house. Natural kinds of things, though, have something molecular about them that allows them to maintain their classifications in a variety of physical states, and their origins are often dealt with theoretically. The origins of humans, for instance, have been explained by various theories.

Natural kinds and artifacts appear to have another dissimilarity especially important here. They exhibit differential developmental patterns. Some investigators, however, presented research that at first inspection suggested that natural kind information might be just as useful to children as it is to adults.

Gelman and Markman (1987) found that when making categorical judgments even young children, if prompted, depended upon the sorts of unseen properties characteristic of natural categories. Children drew more inferences from pictures of things belonging to the same category than they did from things perceptually similar but belonging to different basic level categories (for example, the leaf bug and the leaf). Of the three-year-olds 67% drew inferences from things belonging to the same category as opposed to 36% who drew inferences from perceptually similar things belonging to different categories [$t(15) = 4.35$, $p < .002$].

In a later study (Gelman & Markman, 1991) provided

children with a target picture and two comparison pictures. A pig, for example was shown with a piggy bank and a cow. After discussing both internal and external differences, both three and four year old children selected the correct category match more often than would be expected by chance ($p < .0001$).

Markman (1989) reported that in a study by Gelman and Markman (1987) the perceptual similarity of shark and dolphin was contrasted with the categories of water-breathing and air-breathing. Children used this natural category information 68% of the time which is significantly more than would be expected based on chance ($p < .001$).

Barton and Komatsu (1989, p.438) asked college students to evaluate changes such as the following: Would a pencil maintain its' identity if it looked exactly like a pencil but could not be used to write with? From their research it was concluded that molecular changes were more significant for natural kinds than artifacts ($p < .01$).

These studies could be offered in support of the classical conception of category, for "natural kind" information appeared to be just as useful to children as it was to adults. However, in the Gelman and Markman (1986, 1987) studies and in the Barton and Komatsu (1989) study the categories were identified for the children. This prompting seemed to mask a marked developmental shift. It produced an overestimate of the importance of unforeseen properties to

children.

When Keil (1989) studied natural kinds with no priming category labels given, the results revealed a developmental pattern. When very young children were given an object, say a raccoon, and told that it had been made to look like a skunk and had an odor bag sewn such that it could function like a skunk, they classified it as a skunk. Older children (fourth grade) held that it was still a raccoon.

Young children are certainly predisposed to categorize (Markman, 1989). However, their categorizations appear to be based upon the assumption that a reference is to an entire object rather than the properties of an object (Gelman & Markman, 1985). When a few components of the thing changes, so does the thing itself. In this sense, all things are like artifacts for young children. The most stable categories (natural kinds), however, eventually emerge. It follows that this stability allows them to become richly stocked with both exemplars and properties of exemplars. Hence, they become most useful from the standpoint of inferencing to relatively underdeveloped categories (Mervis, 1987). They are anchors, so to speak, in a sea of instability. This developmental pattern is not predicted by the classical model.

Subordinate, Basic, and Superordinate Category Levels

Rosch (1978) distinguished between three levels of

classification, based on inclusion and abstraction.

1. The Subordinate Level. This is the level at which specific names are given to things (for example, pine tree or Bayer Aspirin).
2. The Basic Level. This is the level at which common or generic names are given to things (for example, tree or aspirin).
3. The Superordinate Level. This is the level at which names are given to collections of things (for example, forest or analgesic).

Basic level categories consist of things that arouse the senses. These are things that can be touched, tasted, smelled, seen, and/or heard, and, hence, directly experienced. Mervis (1987) observed developmental changes both within the level and between the level of the categories used. The following research supports this with a single qualification. While category names according to Green (1989) act as effective cues, the literature suggests that for experts, the subordinate level may elicit as many responses as the basic level (Tanaka & Taylor, 1991). However, children are more often novices than experts.

Rosch (1978) found that children sorted picture cards in a way that indicated that basic level objects were first named. From age three children can sort by basic level almost perfectly. However, three-year-olds were only 55% correct when sorting by superordinate categories. By age

four they were 96% correct when sorting by superordinate categories.

Fenson, Cameron, and Kennedy (1988) used 25 two-year-olds in a study of perceptual matching. Following a brief training period, these subjects matched basic level, superordinate level, and subordinate level objects from the same category. Though the reward paradigm of the study may have imposed artificial constraints, the findings were consistent with those of Rosch (1978). Children chose correct matches 74% of the time ($p < .001$) where perceptual match was very close (same configuration and same basic level category, for example a dark feathered duck and a light feathered duck. The authors report that a chance score would be 59%. When perceptual match was low, for example with a duck and a panda bear, correct choices fell dramatically (36%), even though a duck and a panda bear both belong to the same superordinate level category (animal). Only 2 of the 25 children performed above chance levels.

The ability to distinguish and use these three levels of categories, then, develops over time (Keil & Kelly, 1987). The specific age at which this ability could be seen as complete and finished may be variable, however. Anglin (1970) evaluated sorting tasks presented by Bruner and Oliver (1963) and determined that a shift occurs over time from concrete to abstract. Later, Anglin (1977) found that, indeed, basic level categorization was accomplished

correctly by three year old children. But a group of five year old children and a group of adults were also examined. A developmental shift was observed over time from the correct identification of basic level category exemplars toward those of superordinate and subordinate categories.

While Mervis' (1987) work confirmed this basic level preschool ability, Dale (1972) observed that, consistent with Piagetian theory, changes occur in language ability around age six or seven (during the later preoperational or early concrete operational periods), and reliance on perceptual properties lessens (necessary for basic level categorizations). The child now begins to be able to use what Dale called "symbolic" representations. Older children, Dale observed, are more likely to use superordinate terms like "tool."

A developmental pattern, then, not predicted by the classical model appears to exist with regard to vertical classifications.

Taxonomic/Thematic

The final developmental shift to be examined here concerns the relationship between categories as much as it does the substance of categories. When given the opportunity or in an unlabeled condition (Markman, 1987; Markman & Hutchinson, 1984), young children preferred to sort on the basis of theme as opposed to taxonomy. For

example, when given a list of words including dog, cat, and bone, children grouped dog with bone (when no category label was provided). This is clearly not predicted by the classical model.

Fivush (1987) argued convincingly that young children prefer thematic groupings. This accounts for the development of scripts which represent themes (or events). Taxonomic categorizations based upon listings of either exemplars or attributes were not the preferred sorting style until children were much older. These findings are predicted by Kiel's (1987) work and confirmed (with qualifications) by Fenson, Vella, and Kennedy (1989).

Fenson et al. (1989) used a picture matching technique with children and found that older children are more successful grouping thematically or taxonomically. The developmental shift that is suggested is parallel. However, since the ability begins with theme, the development of the ability to group by taxonomy, by contrast, is more pronounced. This might be seen as consistent with the results of research in other areas of category development reviewed in this study.

If the shift toward the development of categories based on taxonomies is the result of the application of multiple rules, then increased rule development over time might be predicted. Ward et al. (1989) examined the number of rules used in categorizations. Subjects were given descriptions

and pictures with nonsense labels. These pictures of similar or dissimilar objects were presented for categorization judgments. Consistently, three sets of materials used with preschoolers indicated that they were significantly more likely to use single attribute rules ($p < .01$, $p < .05$, $p < .01$). The tendency to use single attribute rules was also present in a fourth set of materials but was not significant ($p > .25$). Second graders and college students, in contrast, indicated dual and triple attribute rules for categorizations. Here the four sets of materials used were significant in all cases ($p < .01$, $p < .05$, $p < .05$, $p < .01$).

Anglin (1970) observed a syntagmatic-paradigmatic shift in the words children use in free association between the ages of five and ten. This shift is similar to the thematic-taxonomic shift. Also, Anglin (1970) directed children to sort words into piles any way they liked. He observed that they often seemed to group the words by theme.

According to the classical model of category-as-container, then, the use of taxonomies based upon features should be developmentally constant. It does not predict a developmental increase in the use of this structure.

An Alternative Explanation of Category

The above challenges indicate that the classical model of category cannot account for the results of much research.

A new model that accounts for qualitative change is needed. Medin and Wattenmaker (1987) suggest such a model.

Medin and Wattenmaker (1987) held that many of the assertions regarding category structure and function do not adequately describe or explain the category. Similarity is insufficient to explain category cohesiveness because no constraints have been identified that will determine which features count as attributes. So perhaps, it was argued, correlations of things account for categories. The claim that attributes, however, are correlated is inadequate because there are so many possible correlations. It is not clear how the correct ones could be selected. In other words, again no useful constraints can be offered.

The process of categorization surely includes relational components. Given a pile of bones, 100 or so pounds of living tissue, a handful of internal organs necessary for self-sustaining acts, and a brain, one would not claim to have a human being. It is the relationship of component parts one to another that results in classification "human." Accordingly, functional relationships must be important to categorization (Medin & Barsalou, 1987).

While that seems true, it is still insufficient to account for categorizations. If functional relationships were all that were important, a motorcycle would share the same basic level category as an automobile because they both

convert a carbon based form of energy into motion which is used to move people from one place to another.

Family resemblance implies relationships and was proposed by Wittgenstein (1953) as one account of category structure. It was suggested that family resemblance correctly described the intercategory relationship for some categories.

Interestingly, however, in a study of University of Illinois undergraduates Medin, Wattenmaker, and Hampson (1987) found that none of the subjects sorted according to principles of family resemblance unless relationships were first explained. These explanations appeared to be parts of theories held regarding the world. This resulted in the conclusion (Medin et al. p. 277, 1987) that:

. . . concepts should be viewed as embedded in theories and are coherent to the extent they fit people's background knowledge or naive theories about the world. Furthermore, the developmental research reviewed by Murphy and Medin (1985) supports the claim that theories help structure categories in even very young children.

While Medin and Wattenmaker (1987) did not specify how these naive theories (which may be thought of as any theory, the objective truth of which is unimportant, that attempts to rationalize the world in some way) come to be instantiated, Barsalou (1987) suggested that there are parallels between this development and the expert/novice distinction. It suffices for the purposes of this discussion, then, to refer to naive theories as those

specified relationships that account for categorization. . . even in children.

Medin and Wattenmaker (1987) and others (McCauley, 1987) suggested that categories are determined by each person's naive theories about the world. It is at this level that the individual, perceptually actively in Gibson's (1966) sense, assigns meaning to an object or activity. It is here that the final definitive act of categorization may occur. It is here that idiosyncracies and age related patterns develop. Whatever organization a finished category may have (family resemblance, prototype, ad hoc, even classical), it is the result of this assessment of the meaning of things in the world. It is suggested by the literature (Langacker, 1987; McCauley, 1987; Medin & Wattenmaker, 1987; 1989; Medin et al. 1987), then, that concepts should be viewed as process driven and embedded in naive theories of relationships, and, above all, developmental.

Reading and Categories

It has been shown that categories change in qualitative ways. One theory which will account for this has been briefly suggested. It is plausible to suggest that an understanding of category development is critical to an understanding of reading development, for the relationship between the two appears to be a very close one.

Cognitive explanations of the reading process consist of an understanding of the whole as an entity with a number of parts which might be assembled or reassembled in a variety of ways (Goodman, 1985; Langacker, 1987; Rumelhart, 1985). These sorts of models are highly interactive and depend upon some organization of knowledge (Adams and Collins, 1985). Recent literature suggests that this knowledge is developmental and organized on some basis of categorized (Rudzka-Ostyn, 1989) information array (Medin, 1986).

The developmental patterns observed in categories offer suitable explanations for various findings in the area of reading. Marks, Doctorow, and Witrock (1974), for example, found that comprehension and retention were greatly increased by the use of more familiar terms rather than less familiar terms from the same category. Krashen (1985) found that children learn to read as well if not better by simply reading. Goodman (1992) supports this. Krashen's results may have an explanation that involves the interaction of organized information (Carnine, 1990). McKeown, Beck, Sinatra, and Loxterman (1992) examined the effects of well-developed background knowledge on different forms of text. Their findings suggested category mapping. Words are symbols used to represent categories (Anglin, 1977; Markman, 1987). As such they become a method of summarization, of enhanced cognitive efficiency (Matlin, 1989). From what has

been stated it can be asserted that the meaning of words, rules for the production of meaning, the very code itself must be known by the reader for comprehension to result. This information is stored in categories which, during the reading act, must map more or less completely onto the categories intended by the author (Carnine, 1990).

The literature offered above suggests that until the categories themselves assume finished proportions, very little accurate mapping of categories onto words in text could occur. This development could be seen to be related to the "stages" of development identified by many authors (for example, Chall, 1967).

Adams (1980) held that reading involves knowing, which may be thought to include robust categories of information. Printed material, then, is only a code for cuing information already possessed by the reader. It follows that meaning is in the author and the reader (Goodman, 1985). Meaning is the desired outcome of the reading act. It is gained, Gibson (1985) noted, by means of a process that involves the increasing specificity of correspondence between stimulation and discrimination, the optimization of attention, and the increasing economy in the gaining of information from print. In a general sense, this seems to correspond with the above observations about the development and mapping of categories of information onto words in text, and with what was suggested by the literature about information stored in and

retrieved from categories.

The classical position suggested that categories might be formed on the basis of similarity (Medin, 1989; Medin & Wattenmaker, 1987) or dissimilarity (Langacker, 1987; Markman, 1987, 1989). For this conception of category to be correct, however, a relatively static knowledge base (Barsalou, 1987) with absolute boundaries would have to be present. Also, it follows from the classical conception of category that information originates in things external to the mind. This view of category strongly resembles "bottom-up" theories of reading (for example, Gough, 1985a).

The preceding review of the current literature indicated that categories are not formed on the basis of similarity alone; that categories are not static; that the boundaries of categories are not permanently fixed; that the organization of categories appears to follow a developmental pattern. Categories appear to change qualitatively (Anglin, 1977) and quantitatively in rather predictable ways; and this appears to occur cross-culturally (Kiel, 1987).

Piaget (in Flugel, 1964; in Waller, 1977) suggested a developmental pattern not unlike that suggested by the literature on category development. During the shift from the later preoperational period to that of concrete operations, the child makes increased use of information that is not centered on the self. The child moves from a world that depends heavily on perception to a world that

depends heavily upon conception (Dinkmeyer, 1965; Downing and Leong, 1982).

This study asks whether, in fact, category development is related to reading development as suggested by both theory and research as suggested by the literature.

Summary of Major Findings

This review of the literature indicated that the classical conception of category, which used the "container" metaphor, cannot adequately account for the structure and functioning of categories. Rather, the literature seemed to suggest a dynamic developmental conception of category, changing with naive theories about the world. It further suggested that category development may parallel reading development, and, therefore, deserves the attention of reading researchers.

Rosch (1973, 1978) offered an alternative to the classical description of category (in which the category has a central value or prototype and indistinct boundaries) and found the usefulness of these prototypes to be related to age. Kiel (1987) inspected categories of natural kinds and found a definite pattern of development. Rosch (1987) distinguished between levels of category abstraction and inspected the development of these patterns with regard to age. Fivush (1987) described a developmental pattern in the use of taxonomies.

It follows from this that categories will appear to have theory-dependent differential structures and differential functions and that the development of these structures and functions is age related. The literature indicated that at least five aspects of categorization change over time: (1) the use of definitive features, (2) the use of prototypes, (3) the use of natural kinds and artifacts, (4) the use of subordinate, basic, and superordinate levels, (5) the use of taxonomies. It also follows from this that categorization is central to meaning, and, as such should be investigated as a potentially important developmental component necessary for reading.

CHAPTER III

RESEARCH PROCEDURES AND MATERIALS

Introduction

The purpose of this study was to examine the relationship between the development of categories of information and both reading performance on one standardized test of reading and grade level. Developed categories of information appear to be necessary for meaningful reading. Five areas of development have been identified. These concern defining features, prototypes, natural kinds, levels of abstraction, and taxonomies. Generally, the procedures and materials utilized in this study were designed to gather information about these five areas of development of categories and about performance on one standardized reading test. Specifically, the procedures and materials were designed to answer the following questions about the development of five aspects of categorization and about performance on one standardized test of reading:

- I. What is the relationship between the identification of defining features and reading test performance or grade level?
- II. What is the relationship between the production of prototypes and reading test performance or grade level?

- III. What is the relationship between the identification of natural kinds and reading test performance or grade level?
- IV. What is the relationship between the identification of superordinate level categories and reading test performance or grade level?
- V. What is the relationship between the production of taxonomic descriptions of category exemplars and reading test performance or grade level?

The Sample

Because of the reasons Barsalou (1987) cites for the instability of concepts and because of the generally accepted notion that development is somewhat idiosyncratic (Keil & Kelly, 1987), specifying the precise age at which development is completed is not possible. The literature, however, suggested marked category development between grades one and three (ages six and nine). A developmental shift, then, should be apparent between these grade levels. Therefore, first and third graders were targeted for this study.

Two rural east Tennessee schools were selected for participation. More than half of the children of one school were enrolled in the free or reduced lunch program. This information was not available for the other school. In general, however, the economic status of the community was

described by the teachers as average to below average. Both schools were predominately Caucasian. All participants in the study were Caucasian. Written permission was secured from the principals and from the district superintendent in the spring of 1991 (see Appendix A). Letters of explanation and parent permission forms were then sent to the parents or care-givers of the 141 children enrolled in the first and third grades (see Appendix A). The letter informed them of the identity of the author, the purpose of the study, and of the five dollar inducement offered to the school for each completed test and task. One week later a reminder was sent to all of the parents or care-givers.

A total of 89 signed permission slips were returned. The subjects who participated in the initial screening, then, were selected on the basis of parental response rather than by a random selection process.

The Gates-MacGinitie Reading Tests (1989), Level 2, Form K was administered to the children. This was used as an initial screening device. To conform to the age parameters (six through nine) of this study, one ten-year-old child was eliminated. The remaining subjects who produced scores in the bottom three stanines (one, two, and three) and the top three stanines (seven, eight, and nine) were selected for the categorization sorting tasks. A total of 33 children (16 first grade children and 17 third grade children) yielded scores in the top or bottom three stanines

and were given five categorization tasks (see Appendix B).

The first grade sample consisted of seven males. Four were in the low stanine group, and three were in the high stanine group. The first grade sample also consisted of nine females. Three were in the low stanine group, and six were in the high stanine group.

The third grade sample consisted of eight males. Six were in the low stanine group; two were in the high stanine group. The third grade sample also consisted of nine females. Four were in the low stanine group, and five were in the high stanine group.

Data Gathering Procedure

All data was collected between 11:00 A.M. and 2:30 P.M. during a two week period in May, 1991. The initial screening was given during two sessions. Prior to testing, the investigator chatted informally with the groups of children. The children were informed that no grades would be given. Therefore, no grades would be reported to the principal, the teachers, or the parents. Their performance on the test, then, would have no affect on their grades at school. The purpose of the study was explained, and the children were told that anyone who did not wish to take the test was free to return to his or her classroom. No child elected to return to his or her classroom. Personal data, which included name, age, grade level, and sex, was then

taken.

The vocabulary section and the comprehension sections of the Gates-MacGinitie Reading Tests (1989), Level 2, Form K were administered during separate sessions (total testing time was about one hour) for each grade level. This was done to accommodate both the teachers and the children. Absent children were tested on alternative dates. The same form of the test was used with both first and third grade children. Out-of-level norms were used for scoring. Therefore, any differences found between grade level groupings could not be attributed to differences in forms.

This test was used: (1) as an initial screening device to determine reading stanine level, and (2) as a basis to determine the correlation between standardized reading test performance and categorization performance. Specifically, the scores from the Gates-MacGinitie Reading Tests (1989), Level 2, Form K were used to determine the following:

1. Representative data for different subgroups.
2. Individual stanine levels.
3. Correlations between reading test performance and categorization ability.

Those children scoring in the bottom three stanines and the top three stanines were selected for the categorization tasks. Each child was interviewed prior to performing these tasks. A brief informal conversation was conducted to put the child at ease. During this conversation personal data

provided during the initial screening was rechecked. Name, age, grade level, and sex were validated. Children were told that there were no right or wrong answers, and grades would not be assigned. There was, in short, no way to pass or fail the sorting tasks. Moreover, the answers given could in no way impact their grades at school. The children were also told that they did not have to do the sorting tasks. If they wished, they were free to leave at any time. All 33 children cooperated fully.

While the determination of what constitutes a relaxed and cooperative atmosphere is purely subjective, there are some criteria which may be used. Among these are overt anxiousness to end the testing session, general body movements, and length of response.

At the conclusion of the sorting tasks, the children did not seem anxious to return to the classroom. No unusual movements were observed (fidgeting, nail biting, etc.). The length of oral answers was also noted. The answers were not hurried. All the children appeared at ease. All appeared to have done their best. The level of cooperation achieved seemed sufficient to assure a measure of validity.

The individual testing was done in currently unused school classrooms. No time limit was placed on the children. As a result, lengths of the sorting task sessions varied considerably from child to child. The shortest response period was approximately 20 minutes. The longest

response period was approximately 50 minutes.

The Gates-MacGinitie Reading Tests (1989), Level 2, Form K is a widely used non-diagnostic screening device, consisting of a vocabulary and a comprehension section. The vocabulary section of the test included 45 pictures with 4 choices of single word identifications. The children were instructed to select the word that best described the picture. These selections were of varying difficulty and involved the use of nouns, verbs, adjectives, and adverbs. Some early readers may, in fact, be unable to read silently. If the words in the test were read orally, they would involve the pronunciation of multisyllabic words, silent letters, long vowels, short vowels, r-controlled vowels, digraphs, and diphthongs.

The comprehension section of the test included 46 short reading passages. Above the passages were three pictures. The children were instructed to select the picture that best described the passage. These passages were of varying difficulty and varying subject matter and included from one to six sentences. The sentences were simple, compound, and complex and included both declarative and interrogative types. The comprehension passages principally consist of the identification of main idea, sequence, cause and effect, and contrast and comparison.

The Gates-MacGinitie Reading Tests (1989), Level 2, Form K was normed on a population of approximately 42,000,

representing 67 school systems. These were stratified by region, size, and socioeconomic characteristics. The equating sample consisted of 25,000 students selected from 36 school systems. The test is reported to have high reliability coefficients. It also provided out-of-level norms which were particularly useful for this study. The same test could be administered, then, to both first and third grade children. The same pictures from the vocabulary section could also be used with both groups, thus increasing confidence in the compared scores.

While all but the Readiness levels have two forms, only Level 2, Form K was required. Similarly, though the test can provide percentile ranks, normal curve equivalents, grade equivalents, scaled scores, and stanines, only stanines were required.

The Gates-MacGinitie Reading Tests (1989), Level 2, Form K consists of two subtests. The vocabulary subtest was given first. The children were given exactly 20 minutes to complete the test. It consisted of 45 questions that involve matching a picture with one of four words that are thought to describe that picture. The questions required decoding skills. The authors of the test held that the central task for the child was to "sound out" the appropriate word.

The comprehension section of the test was given last. The children were given 35 minutes to complete the test. It

consisted of 46 questions of varying lengths. Each question asked that a passage be read and matched with the correct one of three descriptive pictures. Here the authors of the test held that the central task was to indicate an understanding of printed material. The levels of inference required to understand the passage or the questions were not uniform. Neither were the length and complexity of the printed passages uniform.

Sorting Task Development

Sorting and picture identification tasks have often been effectively used for the identification and measurement of categorization ability and performance (Carabine, 1991; Fenson, Vella, & Kennedy, 1989; Markman, 1987, 1989; Markman & Hutchinson, 1984; Rosch, 1987; Tversky, 1985).

Therefore, sorting tasks were designed to test the state of development of certain aspects of categorization ability known to change over time. Pictures selected from the vocabulary section of the Gates-MacGinitie Reading Tests (1989), Level 2, Form K were used (with the permission of the publisher - see Appendix A) for the sorting tasks.

The original sorting tasks were designed to measure the development of certain characteristics of category usage. A pilot study, however, conducted during January and February, 1991 at a regional private school suggested refinements in the initial sorting task design. The

following changes were made during these sessions:

1. The ordering of the questions appeared to be a factor in sustaining interest. Those questions that seemed to best familiarize the child with the sorting tasks and that seemed to produce the longest response times were given first.
2. During the pilot study, many pictures were used. The responses generated during these sessions suggested that some pictures were either unclear or misleading. These pictures were dropped from the finished sorting tasks.
3. The directions used for the study were similarly refined during the pilot sessions. The finished directions appeared to have been clear, meaningful, and brief.
4. Some alterations in the pictures were found to be necessary. A picture of a round table, for example, was used. Several other pictures of round objects were also used to suggest a category of things that are round. The picture of the table, however, was incomplete. Therefore, the line necessary to complete the circular table top was drawn in. Photocopies were made so that the picture appeared as a complete representation, no different qualitatively from the other pictures. It was also necessary to eliminate information

from those pictures used to evaluate natural kind sorting ability.

The finished questions including the directions given to the children appear in Appendix B. They are considered to reflect the concepts described in the preceding material.

The Sorting Tasks

It was previously indicated that the literature supports the use of sorting tasks in the evaluation of categorizations. Cited authors (Carabine, 1991; Fenson, Vella, & Kennedy, 1989; Markman, 1987, 1989; Markman & Hutchinson, 1984; Rosch, 1987; Tversky, 1985) have successfully used sortings in the study of categorization. Hence, sorting tasks were used to answer the five questions.

Statistical techniques appropriate to the question asked and the sorting task design were utilized. The calculations of means, medians, standard deviations, and variances were used to describe samples. The Pearson Product Moment coefficient of correlation and attendant critical values, probability tests, and the t test of the means were used for sample comparisons.

The sorting tasks were designed to measure characteristics of categorization known to change over time. The pictures used were selected from the vocabulary section of the Gates-MacGinitie Reading Tests (1989), Level 2, Form K. The Gates-MacGinitie Reading Tests have been in use for

many years. It was, therefore, assumed that the pictures were valid representations of the named objects or situations. The pictures were cut out for individual presentation and manipulation. They were laminated for easy and clean repeated use. In some cases (where noted) they were altered so that they would be task appropriate. The following is a description, in order, of the administration of the sorting tasks:

Prototype Sorting Task. Each child was seated at a table and told that he/she would shortly be given some pictures. The child was also told that he/she would be asked to sort them into piles of things that are alike. Before administering the test, however, each child was provided with an example. Ten pictures were shown. They included (1) a child with birds, (2) people eating, (3) a child on a balance bar, (4) children playing, (5) a child playing with a toy airplane, (6) a one-way sign, (7) a theater marquee, (8) a party invitation, (9) a car raised on a jack, (10) a dump truck. These were unused pictures also selected from the Gates-MacGinitie Reading Tests (1989) Level 2, Form K. Each child was shown that these pictures could be divided into two piles of pictures, those with people and those without people. Each child was then shown that the pictures could also have been divided into three piles, those with signs or notices, those with children

playing, those that were remaining. It was pointed out to each child that many other sortings might be possible. Each child was then asked if the procedure was understood. All of the 33 children tested indicated that the procedure was understood.

Following this example, the child was given ten new pictures (poster, coin, candle, saucer, medal, soup, shelter, teakettle, pipe, ring). He/she was instructed to sort them into piles of things that appeared to be alike. Further, the child was told to sort them any way that was wished and to take as long as was needed.

When the sortings were completed, the child was asked to label the kinds of things in each pile. It was decided that more information might later be required, or that the labels given to the piles by the child might be unclear. Hence, after the pile was labelled, two further questions were asked. First the child was asked if there was anything else he/she wished to say about the pile of pictures. Second, the child was asked why he/she decided to call them by the given label.

Once a clear understanding of the category had been established, the child was asked what the best example (prototype) of the category might be. The child's own category name was always used in the phrasing of this question.

This procedure was repeated until all of the piles had

been given a category name and each had a labelled prototype. Since categorization abilities change over time, it was anticipated that not all of the children would select prototypes for all of the categories. Some categories of things that cohered for structural and/or functional reasons, then, did not result in a named prototype. In some cases the children elected to develop residual categories. These consisted of things that seemed to cohere only because they did not belong elsewhere. Frequently these categories did not result in a named prototype.

Defining Feature Sorting Task. Since the children were familiar with the sorting task procedure, lengthy instructions were not required for the remaining sorting tasks. Hence, the children were told that they would be given pictures to look at. The pictures included a court room, a child under a table, a medal, children with a ball, a coin, a ring, a saucer, and a soup bowl. The children were asked in what way these things were alike.

Each picture had a circular or round element. There was no other reasonable unifying property. Any feature besides roundness would trivialize or eliminate the difference between this category and other categories. The defining feature of this group of pictures, then, was roundness.

Natural Kind Sorting Task. Since some things are not easily classified as either natural kind or artifact, it was decided that pictures used for this task would be of living things. These are always natural kinds. The literature indicated that categories of natural kinds are exceptionally rich in information. It follows that exemplars of natural kinds provide redundant information. Hence, a measure of classification ability of natural kinds would be provided by the identification of partially completed pictures. If a child had richly stocked natural kind categories at his/her disposal, minimal information might serve as a sufficient cue to recall the category name.

During the pilot studies, several variations were attempted. It was found that the identification of a living thing did not required an entire picture. It did, however, generally depend upon characteristics of the uppermost part of the body. It was also found that to achieve a distribution of scores, 80% of the picture had to be eliminated.

The results of the pilot study further indicated that the resulting partial pictures were too small to be clearly viewed. Varying enlargements were attempted. At last it was determined that a 50% enlargement of the partial picture could be clearly viewed.

Each child, then, was told that he/she would be given some pictures that would not be complete. Each child was

further instructed to identify the pictures. The partial pictures included a bear, a moth, a cow, a bird, and a clown.

Non-basic Level Sorting Task. Each child was provided with an oral example prior to engaging in this task. Each child was told that if given a picture of a cat, a dog, and a rabbit and asked to provide a single name for them, the result might be the name "animal." Each child was then given three pictures (a table, a table and chairs, and a bench) and asked to give them one name. This task required the production of the superordinate or non-basic level label "furniture." Functional descriptions were considered to be incorrect.

Taxonomic Sorting Task. Taxonomies were deemed to be categorizations based upon listings of features and similarities rather than relationships (themes). Hence, for this task, each child was given three pictures one at a time and instructed to describe or tell about the picture. These pictures included a building, a collar, and a bow. Single words, phrases, or clauses were taken to be taxonomic items if they were used as adjectives.

Since categorizations by theme are inversely related to categorizations by taxonomy (see Appendix C), some supporting evidence was provided by an analysis of data

collected for development of ability to produce prototypes. These categorizations were evaluated for the presence of themes or relationships. A category was considered to be organized by theme if one exemplar was described as acting upon another exemplar. To be counted as such, exemplars had to be described in a subject/verb/object sequence, with one exemplar as a subject and one as an object.

Recording Observations

The 33 individual sorting task sessions were recorded in writing. An audio tape recording was made on a high quality voice activated tape recorder. Sony Type II High Bias Tape was used to minimize the background "hiss" found to some degree on all tape recordings. The result was a clear corroborative recording. Following each series of daily sessions, the tape recordings were used to substantiate the written recordings. Each day the tapes were erased and a new series was recorded.

Procedures for Analysis of Data

The procedure for the analysis of data involved within group comparisons and between group comparisons for each of the questions. Hence each question was examined two times. Pattern probabilities were manually calculated. All other data analysis utilized the Minitab (Addison-Wesley, 1987) computer program.

Within Grade Level Treatment of Data

Initially, the means and standard deviations of for the five variables (and stanine scores) were described. To determine the significance of the difference between means, one-tailed t -tests were calculated.

The low readers and the high readers within each grade level were compared. Probabilities of the scores of a given group consistently exceeding the scores of another group were calculated according to the following formula:

$$P = \frac{C}{n} \cdot \frac{p^r}{p} \cdot \frac{q^{n-r}}{q}$$

where P = probability; C = the number of combinations r at a time; n = the number of trials; p = the probable proportion of successes; q = the probable proportion of failures; r = number of successes.

Next, each group was inspected for the relationship between the variables and the stanine scores. These examinations employed the use of the Pearson Product Moment coefficient of correlation. The production or identification of prototypes, natural kinds, and taxonomic lists of features were compared to stanine scores. The frequencies of identifications for defining features and superordinates were similarly compared. Harshbarger (1971) suggests the use of the Pearson Product Moment coefficient of correlation for use with dichotomous variables such as

these. It was further suggested in the literature that the interpretation of the coefficient of correlation depends upon the subject matter (Gay, 1987; Mueller, Schuessler, & Costner, 1970). The interpretation of the results, then, were based upon Davis, 1971 (in Curd, 1983, p. 84). The scale of correlations suggested by Davis follows. It will be used with the caveat that it is only a rough guide to interpretation.

+ 1.00	A perfect positive relationship
+.70 or higher	A very strong positive association
+.50 to +.69	A substantial positive association
+.30 to +.49	A moderate positive association
+.10 to +.29	A low positive association
+.01 to +.09	A negligible positive association
.00	No association
-.01 to -.09	A negligible negative association
-.10 to -.29	A low negative association
-.30 to -.49	A moderate negative association
-.50 to -.69	A substantial negative association
-.70 or lower	A very strong negative association
- 1.00	A perfect negative relationship

While the relatively small sample sizes prevented conclusive results, any consistency of the results was considered suggestive.

Between Grade Levels Treatment of Data

Comparisons between the groups were made initially by inspecting group means and standard deviations. Next, the probability of the score of one group consistently exceeding the score of another was calculated for each of the variables using the formula given above.

Summary

All 141 first and third grade children were given parental permission forms. Eighty-nine signed forms were returned. These children were given the Gates-MacGinitie Reading Tests (1989), Level 2, Form K during May of 1991. Children scoring in the bottom three stanines (stanines one, two, and three) and the top three stanines (stanines seven, eight, and nine) were selected for further study. These 33 children were given five sorting tasks during May of 1991.

Most of the data was analyzed using the Minitab (1989) computer program. It was described in terms of means and standard deviations. A t-test was used to determine the significance of the differences between group means. A probability test was used to evaluate the likelihood of the occurrence of particular patterns, and the Pearson Product Moment coefficient of correlation was employed in the analysis of associations.

CHAPTER IV

PRESENTATION OF DATA

Introduction

The purpose of this study was to examine the relationship between category development, as measured by five categorization (or sorting) tasks, and both reading performance, as measured by the Gates-MacGinitie Reading Tests (1989), Level 2, Form K, and grade level. Five questions were generated to guide the inquiry:

- I. What is the relationship between the identification of defining features and reading test performance or grade level?
- II. What is the relationship between the production of prototypes and reading test performance or grade level?
- III. What is the relationship between the identification of natural kinds and reading test performance or grade level?
- IV. What is the relationship between the identification of superordinate level categories and reading test performance or grade level?
- V. What is the relationship between the production of taxonomic descriptions of category exemplars and reading test performance or grade level?

The results of the study are presented in three sections (see Appendix D for raw data). The first section will provide an overview of data. The second section will focus on an application of the data to each of the questions. The third section is a summary of the chapter.

The first section and the second section include an analysis of means, standard deviations, probability tests, and Pearson Product Moment coefficients of correlations. These sections address the data collected for the prototype scores, natural kind scores, and taxonomy scores. These scores represent incidents of categorizations. The first and second sections also address the data for the defining feature task and the superordinate task. The data for these two tasks represent simple counts of the subjects making correct responses. For all five tasks, t -tests are used to evaluate the means. These t -tests of the means are made between the high reading groups versus the low reading groups. As a result, this presentation is given in positive values.

Grade level comparisons are also offered in two ways. First, comparisons are made between the means of the first and third grade subjects. Second, to maximize the differences, comparisons are made between low first grade readers and high third grade readers.

An analysis of data regrouped by reading level is given. This is presented so that comparisons can be made

between mean differences of the task scores by grade level and by reading level.

The final section of Chapter IV was a summary of the results.

Overview of the Data

First Grade Low Readers

The first grade low readers achieved a mean stanine score of 2.29 on the Gates-MacGinite Reading Tests, 1989, level 2, form K. The standard deviation was 0.49. Table I exhibits the means and standard deviations of the sorting

Table I

Comparison of First Grade Low Reading Group (FLR) and
First Grade High Reading Group (FHR)

Group	N	Means	SD	t-test	df
FLR defining features	7	0.00	0.00	---	---
FHR defining features	9	0.56	0.53		
FLR prototypes	7	2.00	1.41	0.73	14
FHR prototypes	9	2.44	1.01		
FLR natural kinds	7	2.29	1.38	0.80	14
FHR natural kinds	9	2.78	1.09		
FLR superordinates	7	0.00	0.00	---	---
FHR superordinate	9	0.11	0.33		
FLR taxonomies	7	4.43	3.74	0.88	14
FHR taxonomies	9	6.33	4.66		

Levels of Significance: * = .05

task scores and the reading test scores. None of the readers in this group identified defining features or superordinates.

Table II shows the correlation between the performance of the first grade low readers on the categorization (or sorting) tasks and performance on the Gates-MacGinitie Reading Tests (1989), Level 2, Form K. Since the first grade low readers failed to identify any defining features or superordinate categories, no correlations are shown for these two tasks.

Table II

Correlation Between Stanine Score and Categorization Task Performance - First Grade Low Reading Group (FLR) And First Grade High Reading Group (FHR)

Categorization Task	N	r
FLR Defining Features	7	---
FHR Defining Features	9	0.40
FLR Prototypes	7	0.73
FHR Prototypes	9	-0.55
FLR Natural Kinds	7	-0.39
FHR Natural Kinds	9	0.18
FLR Superordinates	7	---
FHR Superordinates	9	0.75*
FLR Taxonomies	7	-0.08
FHR Taxonomies	9	-0.25

Levels of Significance: * = .05.

First Grade High Readers

The mean reading stanine score for this group on the Gates-MacGinitie Reading Tests (1989) Level 2, form K, was 7.56, and the standard deviation was 0.73. The means and standard deviations calculated for the first grade high readers on the categorization are presented in Table I. The correlations between categorization tasks and reading stanine scores are given in Table II.

First Grade High Readers Versus First Grade Low Readers

The first grade high readers produced higher mean scores on all of the sorting tasks. The formula offered in the previous chapter was used to calculate the probability of the occurrence of this pattern (five high scores out of five possible high scores). It was found that the probability of the occurrence of this pattern was .03, where the probability of each success is .50 and of each failure is .50.

Since the first grade low readers failed to produce scores for the defining features task or the superordinate task, tests of significance could only be used to test three sets of means. Hence, *t*-tests were used to test the means of the prototype tasks, the natural kind tasks, and the taxonomy tasks for the first grade high readers and the first grade low readers. Tests of significance of the means for first grade low readers and first grade high readers are

presented in Table I.

All of the scores of the first grade high reading group were higher on the categorization tasks than the scores of the first grade low reading group. None of differences between the means, however, reached the .05 level of significance.

Third Grade Low Readers

The mean reading stanine score on the Gates-MacGinitie Reading Tests, 1989, level 2, form K, was 2.30. The standard deviation was 0.95. The means and the standard

Table III

Comparison of Third Grade Low Reading Group (TLR) and Third Grade High Reading Group (THR)

Group	N	Means	SD	t-test	df
TLR defining features	10	0.20	0.42	t = 1.60	15
THR defining features	7	0.57	0.54		
TLR prototypes	10	1.60	1.27	t = 1.70*	15
THR prototypes	7	2.57	0.98		
TLR natural kinds	10	3.10	1.10	t = 1.60	15
THR natural kinds	7	3.86	0.69		
TLR superordinates	10	0.20	0.42	t = 0.99	15
THR superordinates	7	0.43	0.54		
TLR taxonomies	10	8.20	5.22	t = 0.94	15
THR taxonomies	7	11.00	7.12		

Levels of Significance: * = .05

deviations for the categorization tasks were also calculated for the third grade low readers. This data is presented in Table III.

Table IV shows the correlation between the categorization scores of the third grade low readers and stanine scores on the Gates-MacGinitie Reading Tests (1989), level 2, form K.

Table IV

Correlation Between Stanine Score and Categorization Task Performance - Third Grade Low Reading Group (TLR) And Third Grade High Reading Group (THR)

Categorization Task	N	r
TLR Defining Feature Task	10	-.17
THR Defining Feature Task	7	.75*
TLR Prototype Task	10	-.17
THR Prototype Task	7	.23
TLR Natural Kind Task	10	-.35
THR Natural Kind Task	7	.71*
TLR Superordinate Task	10	-.72*
THR Superordinate Task	7	.42
TLR Taxonomy Task	10	.28
THR Taxonomy Task	7	.66*

Levels of Significance: * = .05.

Third Grade High Readers

The mean reading score for this group, as measured by the Gates-MacGinitie Reading Tests, level 2, form K, was

7.57, with a standard deviation of 0.54. The means and the standard deviations for the categorization tasks for the third grade high readers were also calculated. This data is presented in Table III. Table IV presents the correlations between reading test scores and categorization task scores.

Third Grade Low Readers Versus Third Grade High Readers

The third grade high readers produced higher categorization task scores than the third grade low readers on all of the categorization tasks. The probability formula given above was employed to calculate the likelihood of the occurrence of five higher scores on the five categorization tasks by this group. The probability of the occurrence of this pattern was found to be .03, where the probability of success is .50 and the probability of failure is .50.

Table III presents the tests of the differences between the means.

All First Grade Subjects Versus All Third Grade Subjects

Table V presents a comparison of all first grade readers (both high and low) with all third grade readers (both high and low). The mean for reading stanine (first grade) was 4.93. It was also 4.93 for the third grade. The standard deviations for both groups were identical (3.73). The groups were, then, indistinguishable with regard to grade appropriate reading scores.

Table V shows that in four out of five instances the third grade subjects scored higher than the first grade subjects. The probability of this pattern is .13, where the probability of each success is .50 and of each failure is .50.

Table V
Comparison of All First Grade Subjects (FGS) and
All Third Grade Subjects (TGS)

Group	N	Means	SD	t-test	df
FGS defining feature	16	0.31	0.48	.24	31
TGS defining feature	17	0.35	0.49		
FGS prototype	16	2.25	1.18	.60	31
TGS prototype	17	2.00	1.23		
FGS natural kind	16	2.56	1.21	2.20*	31
TGS natural kind	17	3.41	1.00		
FGS superordinate	16	0.06	0.25	1.75*	31
TGS superordinate	17	0.29	0.47		
FGS taxonomy	16	5.50	4.26	2.11*	31
TGS taxonomy	17	9.35	6.03		

Levels of Significance: * = .05

Table V shows the t-test scores for the mean differences between the first and third grade readers. The differences for three sets of means were significant at the .05 level.

The groups were combined and correlations were

calculated between categorization task score and grade level. Table VI displays these correlations.

Table VI
Correlation Between Grade Level Groups
And Categorization Task Performance

Categorization Task	All 1st & 3rd N	r	Low 1st & High 3rd N	r
Defining Features	33	.04	14	.63*
Prototypes	33	-.11	14	.25
Natural Kinds	33	.37*	14	.61*
Superordinates	33	.30*	14	.52*
Taxonomies	33	.35*	14	.53*

Levels of Significance: * = .05

First Grade Low Readers Versus Third Grade High Readers

A major purpose of this study was to examine the development of categories, especially as this development compares and contrasts with reading development. Performance differences were maximized in this study by comparing the categorizations of low first grade readers with high third grade readers. Data from other tables, then, was regrouped in Table VII.

The third grade high readers produced higher means on all categorization tasks. No test of significance was

Table VII

First Grade Low Readers (FLR) and
Third Grade High Readers (THR)
By Categorization Task Performances

Categorization Task	N	Means	SD	t-test	DF
FLR Defining Features	7	0.00	0.00	----	----
THR Defining Features	7	0.57	0.54		
FLR Prototype	7	2.00	1.41	0.88	12
THR Prototype	7	2.57	0.98		
FLR Natural Kind	7	2.29	1.38	2.69**	12
THR Natural Kind	7	3.86	0.69		
FLR Superordinate	7	0.00	0.00	----	----
THR Superordinate	7	0.43	0.54		
FLR Taxonomy	7	4.43	3.74	2.16*	12
THR Taxonomy	7	11.00	7.12		

Levels of Significance: * = .05; ** = .01

calculated for defining features or superordinates as the first grade low readers failed to identify any instances of either.

Table VI exhibits the correlations between the low first grade readers and the high third grade readers, and categorization task performance. These two groups were conjoined and the correlation was calculated between grade level and categorization task. This resulted in correlations between grade level and performance on the four of the five tasks that were significant at the .05 level.

Three Comparisons of Differences Between Means

The data was regrouped for comparison of the differences between mean scores in Table VIII. Both the low first grade readers and the low third grade readers were grouped together. Similarly, the high first grade readers and the high third grade readers were grouped together.

The differences between the means were found for the first and third grade readers, the low and high readers, and the first grade low readers and the third grade high readers. Table VIII, then, allows for comparisons of the effect of grade level and the effect of reading level.

Table VIII

Comparisons of Differences Between Means
For First Grade and Third Grade,
For Low Readers and High Readers,
For Low First Grade Readers and
High Third Grade Readers

Categorization	First & Third	Low & High	Low First & High Third
Defining Features	.04	.44	.57
Prototypes	.25	.74	.57
Natural Kinds	.85	.49	1.57
Superordinates	.23	.13	.43
Taxonomies	3.85	1.72	6.57

The Questions

Several correlations are detailed in the following section. These are often reported in terms of the Davis, 1971 (in Curd, 1984) scale.

Question I

Question I focused on the relationship between the reading test scores on the Gates-MacGinitie Reading Tests (1989), Level 2, Form K, and the ability to identify defining features. It also focused on the relationship between grade level and the ability to identify defining features.

The defining features identification task involved a dichotomy. Answers were either acceptable or unacceptable. The subjects were shown pictures of a ball, a coin, a ring, a saucer, a bowl of soup, and a medal. The only element shared by all members of this category that could be seen as defining the group was the element of "roundness." Hence, only "round" or synonyms for "round" were acceptable. Features that were not definitive were disallowed. Distinctions that were trivial, for example, "things you can use," were also disallowed.

Table I and Table III show that a larger proportion of children in the higher reading groups in both grades were able to identify defining features. Very similar proportions of the higher reading groups correctly

identified the defining feature - .56 and .57. The standard deviations were also similar - .53 and .54. The magnitude of the differences between the low first grade readers and the high first grade readers, however, was greater than it was for the third grade low readers and the third grade high readers.

The correlations between reading stanine score and defining features are given for grade level one in Table II and for grade level three in Table IV. The higher reading groups produced higher correlations between reading and the identification of defining features than did the lower reading groups. The first grade high readers produced a moderate correlation. The third grade high readers produced a very strong positive association, significant at the .05 level. The correlation for the first grade low readers was not calculated as no identifications were recorded. The correlation between defining features and reading stanine level for the third grade low readers was low and negative.

Table V shows the scores of the first grade readers and the third grade readers. A larger proportion of the third grade subjects identified defining features than did the first grade subjects. The differences between the grade levels, however, were particularly apparent when the low first grade readers were compared to the high third grade readers in Table VII. The low first grade readers failed to identify any defining features, while more than half of the

high third grade readers did so.

The correlations between grade level and the defining feature tasks are provided in Table VI. The association was negligible. However, when the first grade low readers and the third grade high readers were combined and a correlation was calculated, the result was a substantial association significant at the .05 level.

The means of the two high reading groups, then, were higher than the means of the two low reading groups. Stronger correlations between reading stanine levels and defining feature identifications were also evident for both high reading groups. A comparison by grade level indicated that the mean of the third grade subjects was higher than the first grade subjects.

Question II

Question II asked about the relationship between the scores on the Gates-MacGinitie Reading Tests (1989), Level 2, Form K, and the ability to produce prototypes from self-defined categories. It also asked about the relationship between grade levels and the ability to produce prototypes from self-defined categories.

Subjects were given ten pictures and were asked to group them in piles of things that were alike. Next, the subjects were directed to pick the best example of each category that had been created.

Table I shows that the first grade high reading group identified more prototypes than did the low reading group. Table III reveals this same tendency among third grade readers. For the third grade reading groups this difference was significant at the .05 level. The differences between the third grade groups were greater than for the first grade groups. The third grade high readers identified 38% more prototypes than the third grade low readers, while the first grade high readers identified 22% more than the first grade low readers.

The correlations between reading stanine score and the scores on the prototype identification task for the first grade are given in Table II. The two first grade groups showed opposite tendencies. The low reading group scores correlated strongly and positively with reading stanine level. The high reading group scores were also substantial but the correlation was negative.

The third grade correlations for reading stanine score and prototype identification are given in Table IV. The third grade sorting task scores, like those of the first grade, correlated with reading stanine scores negatively for the low readers and positively for the high readers. Both associations were low. The data suggested that the number of prototypes produced might be related to the number of categories produced. However, this alone did not account for the pattern of correlations as the average number of

categories produced by each of the four groups of readers was about 3.00. The differences that did exist were not significant.

The pooled data for the first and third grade is given in Table V. The means for the prototype task were not significantly different. However, when the first grade low readers were compared to the third grade high readers in Table VII, differences did become apparent. The correlation between grade level and prototype production presented in Table VI was low.

The first grade high readers, then, and the third grade high readers produced higher means than did the first grade low readers and the third grade low readers. The correlations between tasks and stanine scores were mixed for this task. The results of between grade level comparisons were also mixed.

Question III

Question III inquired into the relationship between the identification of natural kinds and performance on the Gates-MacGinitie Reading Tests (1989), Level 2, Form K. It also asked about the relationship between the identification of natural kinds and grade level.

The children were given five pictures of living things. The bottom of the picture had been eliminated. The children were asked to identify these pictures.

Table I shows that the mean number of correct natural kind identifications made was greater for the first grade high reading group than for the first grade low reading group. Table III shows this was also true for the third graders. The mean score of the third grade high readers, however, was 25% greater than the third grade low readers while the difference for the first graders was 21%.

The patterns of correlations for the first and third grade reading groups were similar. However, the relationships were stronger for the third grade groups. A moderate negative correlation was found between reading stanine level and natural kind identification for the low first grade readers. A low positive association was found for the high first readers. The third grade low readers produced a moderate negative association while the third grade high readers produced a positive and very strong association, significant at the .05 level.

The comparisons between the first grade readers and the third grade readers are given in Table V. The means for the third graders were higher than the first graders and reached significance at the .05 level. The differences between the means of the first and third grade scores are even greater when the low first readers are compared to the high third readers. The difference between these means were significant at the .01 level. The correlation between grade level and the natural kind task was moderate and significant

at the .05 level. When the scores of the first grade low readers and the third grade high readers were correlated with grade level, the association became substantial and significant at the .05 level.

The comparisons of the means, then, indicated that both the first and the third grade high readers produced higher means than did the first and third grade low readers. The correlations for the low reading groups in both grades were negative while those of the high reading groups were positive. Between level comparisons revealed higher means for the third grade subjects.

Question IV

Question IV addresses the relationship between the development of the ability to identify superordinate groupings and performance on the Gates-MacGinitie Reading Tests (1989) Level 2, Form K. The mean number of subjects identifying superordinate classifications are displayed in Table I for the first grade and Table III for the third grade.

Identification of the superordinate grouping of three pictures was considered indicated if an utterance included the word "furniture." Functional descriptions such as "they are things that you sit on or at" were considered unacceptable. Such descriptions failed to distinguish between the given category and many other possible

categories. After all, the hood of an automobile, a rock, or a pin cushion are also things you could sit on or at; none of which would be expected to be included in the superordinate category "furniture."

The proportion of the group identifying superordinates was greater for the high readers for both first and third grade. The differences between the proportions were greater for third grade. The correlation between performance on the superordinate task and the reading test could not be calculated for the first grade low readers as no identifications were made by that group. However, the correlation for the first grade high readers was .75, which was considered strong and was significant at the .05 level. The scores of the low third grade readers correlated strongly and negatively (significant at the .05 level). The scores of the high readers correlated moderately and positively.

The mean number of identifications for the first and third graders are presented for this task in Table V. More of the third grade subjects correctly identified the superordinate category than did the first grade subjects. The differences in the means of the first grade high and low readers and the third grade high and low readers were not significant at the .05 level. The comparisons between the first grade low readers and the third grade high readers presented in Table VII expand upon the grade level

differences. The first grade low readers all failed to identify a superordinate category. The correlation, however, between grade level and the superordinate identification task was calculated and found to be moderate and significant at the .05 level. When the scores of the low first grade readers and the high third grade readers are correlated with grade level, the result is a substantial association (significant at the .05 level).

The means of the high first and third grade readers were greater than the means of the low first and third grade readers. The correlations between task and stanine level in all measurable cases were moderate to strong. Between grade level comparisons indicated that the third grade correctly identified the group more frequently than the first grade group.

Question V

Question V focused on the ability to produce taxonomies and reading ability, as measured by the Gates-MacGinite Reading Tests, level 2, form K. It also focused on the ability to produce taxonomies and grade level.

Taxonomies were determined to be categorizations based upon listings of features. As such, subjects were given three pictures and asked to describe or tell about them. The number of descriptors in the predicate were tallied. The scores reflect these tallies.

The first and third grade high reading groups produced greater taxonomy scores on the task than did the low reading groups. Tables I and III, however, revealed that the difference was greater for the third grade. In terms of percentages, however, the first grade gain was greater. The first grade high reading group score was 43% higher than the first grade low reading group. The third grade high reading group score was 34% greater than the third grade low reading group.

Further evidence was provided by the categories formed for the prototype task. Categories defined by the action of one exemplar on another exemplar were held to be thematic. The presence of theme was held to be inversely related to the presence of taxonomy. In general, as the number of taxonomic categories increased, the number of thematic categories decrease. This was evident in the correlations between the production of taxonomic and thematic categories for the various grade and reading levels. The correlation between the two kinds of categories was .07 for the first grade low readers and .17 for the first grade high readers. The correlation between the two kinds of categories was -.28 for the third grade low readers and -.37 for the third grade high readers.

The correlations between reading stanine scores and scores for the taxonomy tasks are shown in Table II and Table IV. For both high reading groups the correlations

were greater. However, significance at the .05 level was reached by the third grade high reading group only.

For the combined high and low groups by grade level it was found that the difference between the means of the two groups was significant at the .05 (Table V). The differences between the means of the low first grade readers and the high third grade readers was, as would be expected, also significant at the .05 (Table VII). The correlation between grade level and scores on the taxonomy tasks was significant at the .05 (Table VI).

The means, then, of the high readers in both groups were higher than the means of the low readers. The correlations between task performance and reading stanine level were stronger for the high readers. Between grade level comparisons demonstrated that the third graders produced more taxonomic terms than did the first graders.

Three Comparisons of the Differences Between Means

The task scores for each of the questions varied with reading level and grade level. To combine the effect of these two variables, low first grade readers were compared with high third grade readers in Table VII. The differences between the means were greater here than in any other comparison.

Another way to treat the data, however, was to compare the scores of all low readers and all high readers with all

first graders and all third graders. To compute this the means of all of the low subjects were subtracted from the means of all of the high subjects. The differences between these means are reported in Table VIII. Such means comparisons revealed differences between grade level, reading level, and a combination of the two.

These comparisons indicated that there were greater differences between the means for the defining features task and for the prototypes task when the data were regrouped by reading level. For natural kinds, superordinates, and taxonomies, the differences between the means were greater when the data were regrouped by grade level.

The combined effects are revealed by the means comparisons for the low first grade readers and the high third grade readers.

Summary

Scores from the Gates-MacGinitie Reading Tests (1989), Level 2, Form K, were used to classify first graders and third graders as high readers (stanines 7, 8, and 9) or low readers (stanines 1, 2, and 3). Many of the questions examined in this study were based upon these designations.

Scores on sorting tasks were used to determine category development in five areas. These included the ability to identify or produce defining features, prototypes, natural kinds, superordinates, and taxonomies.

An overview of the data was presented in the first two sections. The results presented in these sections were then applied to each of the questions.

Overview

The mean score and the standard deviation from the Gates-MacGinitie Reading Tests, 1989, Level 2, Form K for the first grade low readers were described. The means and standard deviations were given for performance on each task (Table I). It was found that no first grade low readers identified defining features or superordinates. Correlations (see Table II) between categorization task and reading stanine score were examined. A very strong correlation was found between the ability to produce prototypes and reading stanine score. The remaining correlations ranged from moderate to low. There were no correlations calculated for those tasks where no correct responses were given.

The first grade high readers were described next. The mean score and the standard deviation for the Gates-MacGinitie Reading Tests, 1989, Level 2, Form K were given. The means and standard deviations for task performance for this group were also given (Table I). Correlations between task and reading stanine score were examined. This group produced a very strong association between superordinate identification and reading stanine score (significant at the

.05 level). The correlations for the remaining tasks ranged from substantial to low.

First grade low readers were then compared with first grade high readers. It was found that none of the differences between the means of these groups reached significance. However, the high reading group scored higher on every task. The probability of the occurrence of this pattern was found to be .03.

Data from the third grade reading groups was dealt with similarly. First, the third grade low reading group was described. The mean score and the standard deviation for performance on the Gates-MacGinitie Reading Tests, 1989, Level 2, Form K were given. Means and standard deviations for task performance were provided in Table III. Correlations between reading stanine scores and categorization tasks (see Table IV) were inspected. It was found that one was moderate (natural kinds) and one was very strong (superordinate). All of the remaining correlations were low.

The mean score and the standard deviation for performance on the Gates-MacGinitie Reading Tests, 1989, Level 2, Form K were given for the high third grade reading group. Task performance was indicated in Table III. Correlations were shown to be higher for this group than any of the others (Table IV). Two of the associations were rated as very strong; one was rated as substantial and one

as moderate. Only the association between prototype production and reading stanine score was rated as low.

The means for the low third grade readers and the high third grade readers were compared. The high reading groups scored higher on all of the tasks. The probability of this occurrence was calculated at .03. The difference between prototype scores for the two groups was significant at the .05 level.

The means of all first grade subjects were compared to all third grade subject (Table V). It was found that the third grade subjects scored higher on the sorting tasks four out of five times. The probability of this pattern occurring was found to be .13. For three of the sorting tasks (natural kind, superordinate, taxonomy) the differences between the means were significant at the .05 level.

The correlations between categorization task performance and grade level were examined, and it was found that moderate but significant (at the .05 level) correlations existed between grade level and the natural kinds task, the superordinate level task, and the taxonomies task.

The low first grade readers were compared to the high third grade readers in Table VII. Differences were maximized in this comparison by the combined effects of reading level and grade level.

To further investigate the relationship between grade level or reading level and task performance, the means of all low readers for the sorting tasks were subtracted from the means of all high readers. Similarly, the means of all first graders for the sorting tasks were subtracted from the means of all third graders. Last, the means of the low third grade readers for the sorting tasks were subtracted from the means of the third grade high readers. These differences were compared. The differences for the defining features and prototype tasks were greater for reading level. The differences between the means for tasks utilizing information about natural kinds, superordinates, and taxonomies were greater for grade level.

The Questions

Question I asked about the relationship between reading performance on The Gates-MacGinitie Reading Tests (1989), Level 2, Form K and the identification of defining features.

When within grade level comparisons were made, none of the differences between the means was significant at the .05 level. The means of both high reading groups for both grade levels were higher than the means of the low reading groups for both grade levels. This was also the case for comparisons between the low first grade group and the high third grade group. The high reading groups for both levels also produced stronger correlations than the low reading

groups.

The correlation between the identification of defining features and reading stanine score was very strong and significant at the .05 for the third grade high readers. The correlation between the identification of defining feature and grade level was also strong and significant at the .05 level.

Between grade level comparisons showed that the third grade scores on this categorization task were higher than the first grade scores.

Question II focused on the relationship between prototype production and both reading level and grade level. Scores increased with reading level and grade level. The means of both high reading groups were greater than those of the low groups. The difference between the mean scores for all first and all third grade subjects was significant at the .05 level.

The correlation between reading stanine level and task performance for prototypes was high for the first grade low reading group (.73). The correlation between grade level and prototype task performance was weak.

Question III compared performances on a natural kinds identification task with reading scores and grade levels. Both the high first and third grader readers scored higher on this task than the low first and third grade readers. The difference between the means was significant at the .01

level for the low first grade readers and the high third grade readers, and at the .05 level for all first grade readers and all third grade readers.

The correlation between reading stanine level and performance on the natural kinds task was strong and significant at the .05 level for the high third grade readers only. The correlation between grade level and performance on the natural kinds task was moderate and significant at the .05 level.

Question IV addressed the relationship between performance on the superordinate task and reading level. It also asked about the relationship between performance on this task and grade level. Although none of the mean differences within the two grade levels was significant at the .05 level, the high readers in both the first and third grades more frequently identified the superordinate category than did the low readers in both first and third grade. The difference between means of the first graders and the third graders was significant at the .05 level.

The correlation between task performance and reading stanine level was strongest for the first grade high readers (.75) and significant at the .05 level. The correlation between performance on the superordinate task and grade level was moderate and significant at the .05 level.

Question V inquired into the relationship between the performance on the taxonomy task and both reading level and

grade level. The first grade and the third grade high reading groups scored higher on the categorization task than did the first grade and third grade low reading groups. None of the differences between means within the grade levels, however, was significant at the .05 level. When the means of the low first readers were compared to the means of the high third grade readers, the differences were significant at the .05 level. The difference between the means of all first grade subjects and all third grade subjects was significant at the .05 level.

The correlation between reading stanine score and taxonomy task was only substantial for the high third grade readers. It was significant at the .05 level. It was also significant at the .05 level for grade level and categorization task, though the relationship was not strong.

When the data was presented by reading level (all low subjects and all high subjects), it was shown that the differences between the means were greater for reading level than for grade level in two of the instances.

CHAPTER V

SUMMARY OF RESEARCH

Introduction

The purpose of this study was to examine the relationship between category development, as measured by performance on five categorization tasks and reading level, as measured by the Gates-MacGinitie Reading Tests (1989), Level 2, Form K. Five questions were generated to guide the study. These were:

- I. What is the relationship between the identification of defining features and reading test performance or grade level?
- II. What is the relationship between the production of prototypes and reading test performance or grade level?
- III. What is the relationship between the identification of natural kinds and reading test performance or grade level?
- IV. What is the relationship between the identification of superordinate level categories and reading test performance or grade level?
- V. What is the relationship between the production of taxonomic descriptions of category exemplars and reading test performance or grade level?

Two schools in a rural East Tennessee county agreed to participate in this study. The population for the study included all 141 first and third grade children in these schools. Permission was sought from and granted by the county superintendent of schools, the building principals, the classroom teachers, the parents or care-givers of the children, and the children themselves. Signed forms were distributed and received from the superintendent, the building principals, and the parents (see Appendix A). Verbal consent was received from the teachers and the children.

The Gates-MacGinitie Reading Tests (1989), Level 2, Form K, was administered to the 89 first and third grade children in both schools who returned a signed permission slip, agreed to participate in the study, and fell within the age parameters of the study. Children scoring in stanines one, two, and three were determined to be low readers. Children scoring in stanines seven, eight and nine were determined to be high readers. The total resulting sample included 16 first grade children (6 low readers and 9 high readers) and 17 third grade children (10 low readers and 7 high readers).

Five tasks thought to measure category development were administered to these 33 children (see Appendix B). These tasks measured category performance in the following areas: defining features, prototypes, natural kinds,

superordinates, and taxonomies.

The analysis of the data included the use of means, standard deviations, probability tests, one-tailed t -tests, and Pearson Product Moment coefficients of correlation.

The analysis of data suggested possible relationships between both reading ability and grade level and:

1. The ability to identify defining features.
2. The ability to produce prototypes.
3. The ability to identify natural kinds.
4. The ability to identify superordinates.
5. The ability to produce taxonomies.

The data suggested that in some cases category development seems to have a stronger association with grade level, and in some cases category development appears to have a stronger association with reading level. In addition, the results of this study indicated that the patterns of development of the different components of categories may vary.

Conclusions

Although the data base for this study was very small, it, nonetheless, resulted in the production of both general and specific conclusions. These conclusions all assume that the development of categories is affected to some as yet unknown extent by direct instruction.

General Conclusions

1. Progress in the study of category development and reading requires more precise measurement instruments.
2. The categorized organization of background knowledge may both affect and be affected by reading performance.
3. While categorization task performances relate to reading level and grade level, the strength of the relationship depends upon the aspect of categorization under consideration.
4. Reading test performance may to some degree be a measure of categorization development.

Specific Conclusions

1. Some aspects of categorization may be strengthened through educational experience; some may not.
2. The teaching of vocabulary has often focused on definitive features. This may not be an age-appropriate method of establishing the connection between graphic symbols and meaning.
3. The data suggests that words representing prototypes will frequently carry more information for faster application. For readers, these words will become increasingly useful with age.
4. Category distinctions based upon origins are

difficult for the early reader. Origins should be taught and natural kinds used as rich sources of information.

5. Words representing non-basic level categorizations should often be avoided until developmentally appropriate.
6. Taxonomies probably reflect a developmental pattern that is inversely related to that of thematic development. As such, theme should be used frequently for early readers. It is perhaps the case that taxonomies may be profitably used only with more mature readers.

Discussion

For convenience, the general conclusions and discussion will focus on reading evaluation, while the specific conclusions and discussion will focus on reading instruction.

General Conclusions: Reading Evaluation

The data offered here suggest the consideration of category development in reading evaluation. A test of category development might differentiate children without developed categories from children who cannot read well for other reasons.

While the data presented in the previous chapter were

not conclusive, a pattern was evident. In all cases the high reading groups performed more successfully on the categorization tasks than the low reading groups. In 80% of the reported associations between reading level and categorization task, correlations went from negative for the low reading group to positive for the low reading group or from lower for the low reading group to higher for the high reading group. Four out of five correlations between grade level and sorting task were significant at the .05 level for the low first graders and the high third graders.

A more rigorous measuring device must be developed so that these patterns can be more fully investigated. Such a device may establish and describe the exact nature of the relationship between category development and reading. Until such development, reading test results must be viewed cautiously.

Developed background appears, then, to depend upon more than just the quantity of the knowledge. The categorical structure of that knowledge must also be considered. The data collected for this study indicated that for many early readers, categorical structures have not yet fully developed. The data further imply that the development of categorized background knowledge may, indeed, be closely related to reading development.

This development may not, however, be uniform among the investigated aspects of categorization. Although category

development and reading are probably interactive, it follows, nonetheless, that some children designated "poor readers" may, in fact, have incompletely developed categories of background information. Information cues for these children would not be productive, and most typical forms of remediation might prove ineffective.

Inappropriate designations may be the result of inadequate tests. Test inadequacy here may take one of two unattractive forms. First, reading tests may be as sensitive to categorization ability as they are to reading ability. Unplanned sensitivity would serve no useful purpose, acting only to distort the test results. An alternative to this is that reading tests are not sensitive to the development of categories. The result is a less than adequate measuring device for reading. The data, then, indicated a need to examine reading tests. These tests may not employ categorizations in a developmentally appropriate way. Until such an examination occurs, it is recommended that reading evaluators examine test instruments for vocabulary and passages that violate what is known about category development.

Specific Recommendations: Reading Instruction

At present, the effect of direct instruction on the development of categories is unclear. Further study is needed to determine which aspects of category development

respond to direct instruction.

It appears, therefore, prudent at this point to suggest that there may be little benefit to early readers in the inappropriate use of defining features, non-basic level categories, or taxonomies. However, prototypes and natural kind distinctions may prove to be rich sources of information for early readers.

It follows from this and from the data collected for the study suggest that the use of the following will only be selectively appropriate:

1. Definitive features. Vocabulary exercises utilizing definitions will be of limited value for early readers.
2. Non-basic level categories. Early readers may have difficulty using terms like forest, furniture, or crowd.
3. Thematic relationships may be more useful for early readers than taxonomies. Hence, fiction may be more appropriate than exposition.

Further study may identify methods of determining word appropriateness consistent with what is known about category development. Vocabulary instruction utilizing definitions is inconsistent with the findings of this study. Definitions are, after all, often little more than listings of definitive features. For some categories, defined boundaries may not exist at all. The boundaries for these

categories have been demonstrated to be fuzzy. Similarly, it might be worthwhile for reading teachers to use caution with words that represent non-basic level categories until development is indicated. The meaning of words like "forest" or "furniture" may be unclear to early readers. Taxonomic categorizations have been demonstrated to be developmental. The early use of thematic representations may be more appropriate than the early use of taxonomic representations.

While the data from this study suggest that taxonomies may also be age-inappropriate, the strong pattern of development suggests that children are at some point able to make use of data presented in this fashion.

The data presented here also further indicated that while definitive features, non-basic level terms, and taxonomies might best be used with caution until development is appropriate, prototypes and natural kinds might be usefully employed as information-rich sources of meaning.

It is clear from the literature on prototypes - and supported by the data in this study - that some words are more useful than others. Lists of prototypes might serve as critical instructional tools for teachers of early readers. Interestingly, the reviewed literature indicated that prototypes are easily identified. Yet it may be found that these are not the words commonly thought to be of particularly high utility. Instruction for early readers

might profitably incorporate what is known about the development of the ability to use prototypes.

Inferencing is a critical component of reading (Goodman, 1985; Tanenhaus, 1988). It depends upon background knowledge (Cooper, 1986). Natural kinds are particularly information-rich sources of background knowledge. Words representing natural kinds (for example, living things), frequently interspersed throughout text, might make meaning much more clear to beginning readers and to children who have reading problems.

Recommendations for Further Research

Both the review of the literature and the results of this study contributed to the recommendations that follow.

The first recommendation is that a much larger data base be obtained. In many cases, of course, there is no certainty the same results would be obtained. If the same results were, however, obtained from larger samples, much higher levels of significance would have resulted.

Second, and equally important, the categorization tasks need several additional incidents of the presented concept. Multiple examples of, say, defining features, would generate measurement scales for each variable that go beyond nominal distinctions. It would eliminate some of the risk that a subject could generally perform a task but could not do so in one particular case. Uniform numbers of tasks might

further allow for more accurate comparisons among the various aspects of category development.

Third, the study of categorization certainly has broad implications for reading assessment and reading instruction. The results of this study indicate a need for the production of a test of categorization development. Such a test would assist in the separation of children who are really poor readers from those who have undeveloped categories.

Fourth, research needs to be directed toward the problem of building mature categories. If it is found that category development responds to instruction, remediation could be employed to bring children to the level necessary for reading to successfully occur.

Fifth, the study of categories must inevitably lead to refined reading evaluational and instructional material. A casual inspection of beginning reading texts indicates frequent violations of the concepts presented and supported by this study. A similar inspection of reading evaluational materials suggests that here, too, violations occur.

LIST OF REFERENCES

LIST OF REFERENCES

- Adams, M. J. (1980). Failures to comprehend and levels of processing in reading. In R. J. Spiro, B. C. Bruce, W. F. Brewer (Ed.), Theoretical issues in reading comprehension (pp. 11-31). Hillsdale, NJ: Erlbaum.
- Adams, M. J. & Collins, A. (1985). A schema-theoretic view of reading. In H. Singer and R. B. Ruddell (Eds.), Theoretical models and processes of reading (pp. 404-425). Newark, DE: International Reading Association.
- Alexander, J. E. (1983), ed. Teaching reading (2nd ed). Boston: Little, Brown.
- Alexander, J. E. & Heathington, B. S. (1988). Assessing and correcting classroom reading problems. Glenview, IL: Scott.
- Anglin, J. M. (1970). The growth of word meaning. Cambridge, MA: M.I.T. Press.
- Anglin, J. M. (1977). Word, object, and conceptual development. New York: Norton.
- Barton, M. E. & Komatsu, L. K. (1989). Defining features of natural kinds and artifacts. Journal of Psycholinguistic Research, 18 (5), 433-447.
- Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. Memory & Cognition, 10, 82-93.
- Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11 (4), 629-649.
- Barsalou, L. W. (1986). Are there static representations in long-term memory? Behavioral and Brain Sciences, 9 (4), 651-652.
- Barsalou, L. W. (1987). Intraconcept similarity and its implication for interconcept similarity. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 76-121). New York: Cambridge University Press.

- Benjamin, H. (1939). The saber-tooth curriculum. New York: McGraw-Hill.
- Bruner, J. S. & Oliver, R. R. Development of equivalence transformations in children (1963). In J. C. Wright & J. Kagan (Eds.), Basic cognitive processes in children. Monograph of the Society for Research in Child Development, 28 (2), (Serial No. 86), 125-143.
- Carabine, B. (1991). Fuzzy boundaries and the extension of object-words. Journal of Child Language, 18 (2), 355-372.
- Carnine, D. (1990, January). New research on the brain: Implications for instruction. Kappan, pp. 372-377.
- Chall, J. (1967). Learning to read: The great debate. New York: McGraw-Hill.
- Chomsky, N. (1959). Review of Skinner's verbal behavior. In Language, 35 (1), 26-58.
- Cooper, J. D. (1986). Improving reading comprehension. Boston: Houghton Mifflin.
- Davis, J. A. (1971). Elementary survey analysis. Englewood Cliffs, NJ: Prentice-Hall. In F. S. Curd (1983). Listening Comprehension: The prediction of reading and language arts achievement in lower elementary grades. Unpublished doctoral dissertation, University of Tennessee.
- Dale, P. S. (1972). Language development: Structure and function. Hinsdale, IL: Dryden.
- Dinkmeyer, D. C. (1965) Child development: The emerging self. Englewood Cliffs, NJ: Prentice-Hall.
- Dolch, E. W. (1939). A manual for remedial reading. Champaign, IL: Garrard.
- Downing, J. & Leong, C. K. (1982). Psychology of reading. New York: Macmillan.
- Feldhaus, R. O. & Honeck, R. P. (1989). The conceptual basis of graded categories. Journal of Psycholinguistic Research, 18 (3), 271-288.
- Fenson, L., Cameron, M. S., & Kennedy, M. (1988). Role of perceptual and conceptual similarity in category matching at age two years. Child Development, 59 (4), 897-907.

- Fenson, L., Veella, D. & Kennedy, M. (1989). Children's knowledge of thematic and taxonomic relations at two years of age. Child Development, 60 (4), 911-919.
- Fivush, R. (1987). Scripts and categories: Interrelationships in development. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 234-254). New York: Cambridge University Press.
- Flugel, J. C. (1964). A hundred years of psychology. London: University Paperbacks.
- Gay, L. R. (1987) Educational research: Competencies for analysis and application (3rd ed.). Columbus, OH: Merrill.
- Geeraerts, D. (1989). Introduction: Prospects and problems of prototype theory. Linguistics, 27 (4), 587-611.
- Gelman, S. A., & Markman, E. (1985). Implicit contrast in adjectives vs. nouns: Implications for word-learning in preschoolers. Journal of Child Language, 12, 125-143.
- Gelman, S. A. & Markman, E. (1986). Categories and induction in young children. Cognition, 23, 183-208.
- Gelman, S. A. & Markman, E. (1987). Young children's inductions from natural kinds: The role of categories and appearances. Child Development, 58 (6), 1532-1541.
- Gelman, S. A. & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. Cognition, 38 (3), 213-244.
- Gibson, E. J. (1985). Trends in perceptual development: Implications for the reading process. In H. Singer and R. B. Ruddell (Eds.), Theoretical models and processes of reading (3rd ed.), (pp.144-173). Newark, DE: International Reading Association.
- Gibson, J. J. (1966). The senses considered as perceptual systems. Boston: Houghton.
- Gillet, J.W. & Temple, C. (1990). Understanding reading problems (3rd ed.). Glenview, IL: Scott.
- Goodman, K. S. (1985). Unity in reading. In H. Singer & R. B. Ruddell (Eds.), Theoretical models and processes of reading (3rd ed.), (pp. 813-840). Newark, DE:

International Reading Association.

- Goodman, K. S. (1992). Whole language research: foundations and development. In S. J. Samuals & A. E. Farstrup (Eds.), What research has to say about reading instruction (2nd ed.), (pp. 46-69). Newark, DE: International Reading Association.
- Goodman, Y. M. (1990). Children's knowledge about literacy development: An afterword. In Y. M. Goodman (Ed.), How children construct literacy: Piagetian perspectives, (pp. 115-123). Newark, DE: International Reading Association.
- Gough, P. B. (1985a). One second of reading. In H. Singer & R. B. Ruddell (Eds.), Theoretical models and processes of reading (3rd ed.), (pp. 661-686). Newark, DE: International Reading Association. Reprinted from J. F. Kavanagh & I. G. Mattingly (Eds), 1972, Language by ear and by eye (pp 331-358). Cambridge, MA: MIT Press.
- Gough, P. B. (1985b). One second of reading: Postscript. In H. Singer & R. B. Ruddell, (Eds.), Theoretical models and processes of reading (3rd ed.), (pp. 687-688). Newark, DE: International Reading Association.
- Harnad, S. (1987). Category induction and representation. In S. Harnad (Ed), Categorical perception (pp.535-565). New York: Cambridge University Press.
- Harris, A. J. & Sipay, E. R. (1985). How to increase reading ability: A guide to developmental and remedial methods (8th ed.). White Plains, NY: Longman.
- Harshbarger, T. R. (1971). Introductory statistics: A decision map. New York: MacMillan.
- Keil, F. C. (1987). Conceptual development and category structure. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 175-200). New York: Cambridge University Press.
- Keil, F. C. & Kelly, M. H. (1987). Developmental changes in category structure. In S. Harnad (Ed.), Categorical perception: The groundwork of cognition (pp. 491-510). New York: Cambridge University Press.
- Kelly, M. H., Bock, J. K., & Keil, F. C. (1986). Prototypicality in a linguistic context: Effects on

- sentence structure. Journal of Memory and Language, 25 (1), 59-74.
- Kletzien, S. B. (1991). Strategy use by good and poor comprehenders reading expository text of differing levels. Reading Research Quarterly 24 (1), 67-86.
- Krashen, S. D. (1985). Do we learn to read by reading? The relationship between free reading and reading ability. In D. Tannen (Ed.), Linguistics in context: Connecting observation and understanding (pp. 269-298). Norwood, NJ: Ablex.
- Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind. Chicago, IL: University of Chicago Press.
- Langacker, R. W. (1987). Foundations of cognitive grammar. Stanford, CA: Stanford University Press.
- Markman, E. M. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 255-287). New York: Cambridge University Press.
- Markman, E.M. (1989). Categorization and naming in children: Problems in induction. Cambridge, MA: M.I.T. Press.
- Markman, E. M. & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic vs. thematic relations. Cognitive Psychology, 16, 1-27.
- Marks, C.B., Doctorow, M.J., and Wittrock, M.C. (1974). Word frequency and reading comprehension. Journal of Educational Research, 67, 259-262.
- Matlin, M. W. (1989). Cognition. Chicago, IL: Holt Rinehart and Winston, Inc.
- McCauley, R. N. (1987). The role of theories in a theory of concepts. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 288-309). New York: Cambridge University Press.
- McKeown, M. G., Beck, I. L., Sinatra, G. M., & Loxterman, J. A. (1992). The contribution of prior knowledge and coherent text to comprehension. Reading Research Quarterly, 27 (1), 78-93.

- Medin, D. L. (1986). Comment on "Memory storage and retrieval processes in category learning." Journal of Experimental Psychology: General, 115 (4), 373-381.
- Medin, D. L. (1989). Concepts and conceptual structure. American Psychologist. December, 1469-1481.
- Medin, D. & Barsalou L. W. (1987). Categorization processes and categorical perception. In S. Hernad (Ed.), Categorical perception: The groundwork of cognition (pp. 455-489). New York: Cambridge University Press.
- Medin, D. L. & Wattenmaker, W. D. (1987). Category cohesiveness, theories, and cognitive archeology. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 25-62). New York: Cambridge University Press.
- Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. Cognitive Psychology, 19 (2) 242-279.
- Mervis, C. B. (1987). Child-basic object categories and early lexical development. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 201-233). New York: Cambridge University Press.
- Mervis, C. B. & Pani, J. R. (1980). Acquisition of basic object categories. Cognitive Psychology, 12, 496-522.
- Mueller, J. H., Schuessler, K. F., & Costner, H. L. (1970). Statistical reasoning in sociology. New York: Houghton Mifflin.
- Murphy, G. L. & Medin, D. L. (1985). The role of theories in conceptual coherence. Psychological Review, 92, 289-316.
- Nagy, W. E. (1988). Teaching vocabulary to improve reading comprehension. Newark, DE: International Reading Association.
- Ozman, H. A. & Craver, S. M. (1986). Philosophical foundations of education. Columbus, OH: Merrill.
- Pflaum, S. W. (1986). The development of language and literacy in young children. Columbus, OH: Merrill.
- Posner, M. I. & Keele, S. (1968) On the genesis of abstract

- ideas. Journal of Experimental Psychology, 77, 353-363.
- Roller, C. M. (1990). The interaction of knowledge and structure variables in the processing of expository prose. Reading Research Quarterly, 25 (2), 79-89.
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), Cognitive development and the acquisition of language (pp. 111-144). New York: Harcourt.
- Rosch, E. (1987). Linguistic relativity. Etc., 44 (3), 254- 279.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), Cognition and categorization (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum.
- Rosch, E., Simpson C., & Miller, R. S. (1976). Structural bases of typicality effects. Journal of Experimental Psychology: Human Perception and Performance, 2, 491-502.
- Ross, B. H., Perkins, S. J., & Tenpenny, P. L. (1990). Reminding-based category learning. Cognitive Psychology, 22 (4), 460-492.
- Rudzka-Ostyn, B. (1989). Prototypes, schemas, and cross-category correspondences: The case for ask. Linguistics, 27, 613-661.
- Rumelhart, D. E. (1985). Toward an interactive model of reading. In H. Singer & R. B. Ruddell, Theoretical models and processes of reading (3rd ed.) (pp. 722-750). Newark, DE: International Reading Association.
- Schaefer, R. L. & Anderson, R. B. (1989). Minitab [Computer program]. New York: Addison-Wesley.
- Schank, R. C., Collins, G. C., & Hunter, L. E (1986). Transcending inductive category formation in learning. Behavioral and Brain Sciences 9 (4), 639-653.
- Skinner, B. F. (1953). Science and human behavior. New York: Macmillan.
- Taylor, D. (1989, November). Toward a unified theory of literacy learning and instructional practices. Kappan, pp. 184-193.

- Tanaka, J. W. & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? Cognitive Psychology, 23 (3), 457-482.
- Tanenhaus, M. K. (1988). Psycholinguistics: An overview. In F.J. Newmeyer, Linguistics: The Cambridge survey. Volume III Language: Psychological and biological aspects. New York: Cambridge University Press (pp. 1-37).
- Tversky, B. (1985). The development of taxonomic organization of named and pictured categories. Developmental Psychology, 21, 1111-1119.
- Underwood, J. & Underwood G. (1987). Data organisation and retrieval by children. The British Journal of Educational Psychology. 57 (3), 313-328.
- Waller, T. G. (1977). Think first, read later! Piagetian prerequisites for reading. Newark, DE: International Reading Association.
- Ward, T. B., Vela, E., Peery, M. L., Lewis, S. N., Bauer, N. K., & Klint, K. A. (1989). What makes a vibble a vibble? A developmental study of category generalization. Child Development, 60 (1), 214-224.
- Wittgenstein, L. (1953). Philosophical investigations. New York: MacMillan.

APPENDICES

APPENDIX A
LETTERS OF CONSENT

Explanation of Purpose of Study to Parents

Dear Parent,

My name is Robert Wheeler. I am a doctoral student at the University of Tennessee; my major area of interest is reading.

My dissertation is a study of the development of the categories of information possessed by the child and how this relates to reading test success. It is my hope that research of this kind will lead to better testing materials and more accurate evaluations.

With your approval, first I will administer the Gates-MacGinitie Reading Tests (1989), Level 2 Form K to your child. Second, I will administer a picture sorting task (using the pictures from the vocabulary subtest of the Gates-MacGinitie Reading Tests (1989)). It will take about one hour to give the test and about half an hour to do the sorting tasks. These will be scheduled during regular school hours. In the event your child is absent on the date of testing, alternative dates will be made available.

On a separate form I have asked for your signature authorizing me to collect data from the test and task and to use that data in a dissertation study. The name of your child will remain confidential. Participation is voluntary. Should you decide to allow your child to participate, you have no obligation other than to let me use the collected data. You may withdraw your child from participation in the study at any time without penalty. There is no risk to your child other than that ordinarily encountered in daily life or during the performance of standardized reading tests.

At the conclusion of the sessions, your child's school will be presented with \$5.00 for each completed test and task. If enough children participate, the contribution to the school could be significant. Please consider allowing your child to participate. I would be grateful, and I am confident that the staff of your child's school would also be grateful.

Sincerely,

Robert Wheeler

Informed Consent and Parental Approval Form

My signature below indicates that I give permission for my child to take the Gates-MacGinitie Reading Tests (1989), level 2, form K and for my child to perform a picture sorting task using the pictures in the vocabulary subtest of the Gates-MacGinitie Reading Tests (1989). My signature below also indicates that I give permission to use data collected from the administration of the Gates-MacGinitie Reading Tests (1989) and the sorting task in a study of category development and early reading. I further understand that I can withdraw my child from participation in the study at any time.

I understand that the name of my child will remain confidential, that is, his/her name will never be used in the study. I further understand that at the conclusion of the two required sessions, my child's school will receive a donation of \$5.00.

Child's Name

Signature of Parent

Signature of Researcher

Date

Robert R. Wheeler
163 Hollow Road
Cosby, TN 37722

Dear Principal,*

I am a doctoral student at the University of Tennessee. My major area of interest is reading.

My dissertation is a study of the development of categories of information possessed by the child and how this relates to reading test success. A comparison will be made between the performance of first grade children and third grade children on the Gates-MacGinitie Reading Tests (1989), Level 2, Form K and five picture sorting tasks. The pictures used are from the vocabulary section of the same test and are held to represent word meanings and categorical information. It is my hope that research of this kind will lead to better testing materials and more accurate evaluations so that reading problems can be more successfully remediated.

The Gates-MacGinitie Reading Tests (1989) are widely used by teachers of reading. The test takes about one hour and can be administered to all of the children at one time. The sorting tasks will take about 30 minutes per child and must be done one child at a time. The total time on task for each child, then, will be about one and a half hours.

No child will participate without signed parental approval. At the conclusion of the research, the school will receive a contribution of \$5.00 for each child that has completed both the sorting task and the reading test. It is hoped that many children will participate.

This project has been approved by Dr. J. E. Alexander, my committee chair and head of the Department of Curriculum and Instruction at the University of Tennessee, Knoxville; Dr. M. Johnson (psychologist and attorney), Department of Psychology, University of Tennessee, Knoxville; Dr. D. Dessart (statistician), Department of Curriculum and Instruction, University of Tennessee, Knoxville; Dr. L. Knight (professor, language arts), Department of Curriculum and Instruction, University of Tennessee, Knoxville.

With your permission the project will be completed before the termination of the 1990-1991 school year. Your signature below will indicate your approval.**

School Principal

Thank you and sincerely,

Robert R. Wheeler

*This form was also used for the letter to the superintendent.

**Subsequent signed letters of approval on county school or elementary school letter-head stationary were secured for and are on file with the CRP committee of the University of Tennessee.

Robert R. Wheeler
163 Hollow Road
Cosby, TN 37722

March 5, 1991

The Riverside Publishing Company
8420 Bryn Mawr Avenue
Chicago, IL 60631

Attn: Office of Copyright Affairs

Re: Request for permission to use aspects of
Gates-MacGinitie Reading Tests (1989) Level 2, Form K.

Dear Sirs:

The purpose of this letter is to request permission to use aspects of Gates-MacGinitie Reading Test (1989), Level 2, Form K in a dissertation study investigating the relationship between cognitive categories and reading ability. This study focuses on sorting tasks involving pictures used in the vocabulary subtest of the Gates-MacGinitie Reading Test (1989).

Given that the specific purpose of the study is to determine if a developmental shift occurs in the categorized information of students who are good readers and older reader, permission is hereby sought to include pictures from the vocabulary subtest of the Gates-MacGinitie Reading Test (1989). Please be assured that this request is for one-time only use and not for the inclusion of intact passages. Acknowledgment of the approval and of the property rights of Riverside Publishing Company of adapted material will appear on the material.

Your consideration of this request and your consent to allow the inclusion of information from your test as cited above will be greatly appreciated.

Sincerely,

Robert R. Wheeler
Doctoral Candidate, College of Education, Department of
Curriculum and Instruction, University of Tennessee,
Knoxville

cc: Dr. J. Estill Alexander, Advisor

rrw



April 23, 1991

Mr. Robert R. Wheeler
Rt. 1, Box 102
Cosby, TN 37722

Dear Mr. Wheeler:

If you wish to administer portions of the Gates MacGinitie Reading Tests in your research, you may do so subject only to routine procedures for test use. Our Customer Service Department shall be able to assist you.

I wish you every success in your dissertation research.

Sincerely,

Fredrick L. Finch
Vice President/Editor in Chief

mkk/24
enclosure

APPENDIX B
CATEGORIZATION TASKS

PROTOTYPE SORTING TASK

Directions:

In a few minutes I'm going to give you some pictures. Then I want you to put them in piles of things that are alike to you. Let me show you what I mean. Spread out 10 pictures (child with birds, people eating, child on balance bar, children playing, child with toy airplane, one-way sign, theater marquee, party invitation, car on jack, dump truck) from the GMRT, 1989, level 2, form K not used in other sorting tasks). I will sort these into two piles of pictures. One is a pile of pictures with people; the other is a pile of pictures without people. I could have sorted these into other kinds of piles. (Spread out all 10 pictures again.) This time I think I'll sort the pictures into three piles . . . signs or notices (invitation, theater marquee, "one-way" sign), children playing (child with airplane, children in a circle, girl on balance beam), and other things (dump truck, people eating, child watching birds, car on jack). Can you help me?

Sorting

Response:

Now it's your turn. (Give child 10 pictures - poster, coin, candle, saucer, medal, soup, shelter, tea kettle, pipe, ring.) Sort these pictures into piles of things that are alike. Sort them any way you wish. Take as long as you need.

Labelling

Response:

What kinds of things are in this pile?

(Prompt) What else can you say about them?

(Prompt) Why did you decide to call them that?

(Prompt Prototype) Is one of these the best
(name child's category label) in the pile?

Labelling
Response:

What kinds of things are in this pile?

(Prompt) What else can you say about them?

(Prompt) Why did you decide to call them
that?

(Prompt Prototype) Is one of these the best
name child's category label) in the pile?

Labelling
Response:

What kinds of things are in this pile?

(Prompt) What else can you say about them?

(Prompt) Why did you decide to call them
that?

(Prompt Prototype) Is one of these the best
(name child's category label) in the pile?

Labelling
Response:

What kinds of things are in this pile?

(Prompt) What else can you say about them?

(Prompt) Why did you decide to call them

that?

(Prompt prototype) Is one of these the best
(name child's category label) in the pile?

Labelling
Response:

What kinds of things are in this pile?

(Prompt) What else can you say about them?

(Prompt) Why did you decide to call them
that?

(Prompt prototype) Is one of these the best
(name child's category label) in the pile?

Labelling
Response:

What kinds of things are in this pile?

(Prompt) What else can you say about them?

(Prompt) Why did you decide to call them
that?

(Prompt Prototype) Is one of these the best
(name child's category label) in the pile?

CHARACTERISTIC/DEFINING FEATURE SORT

Directions: Now I'm going to give you some pictures to look at. (Give child pictures of court room, child under table, medal, children with ball, coin, ring, saucer, soup.)

In what way are these things all alike?

(Prompt) Why did you say that?

NATURAL KIND SORTING TASK

Directions: Now I am going to give you some pictures that aren't complete. Let's see if you can tell me what they are pictures of.

Sorting

Responses: (bear)

(moth)

(cow)

(bird)

(clown)

NON-BASIC LEVEL SORTING TASK

Directions: If I showed you a picture of a cat, a dog, and a rabbit, you might say they all have one name. They are all called animals. Here are three pictures. (Give child pictures of table, table and chairs, bench.) What is a name for all three pictures?

TAXONOMIC SORTING TASK

Directions: Now I'm going to give you three pictures . . .
. one at a time. I'd like you to describe
each picture.

(building) Can you describe or tell about
this picture?

(collar) Can you describe or tell about this
picture?

(bow) Can you describe or tell about this
picture?

APPENDIX C

ASSUMPTIONS, LIMITATIONS, DELIMITATIONS

Assumptions

The following assumptions were used in this study:

1. Words represent categories.
2. Pictures in the Gates-MacGinitie Reading Tests (1989), Level 2, Form K represent the same category represented by the correct attendant word.
3. Sorting tasks and verbal responses are indicators of category development.
4. Correct responses to test questions are indicators of reading ability.
5. Reading difficulties unrelated to information categories under investigation will be randomly distributed.
6. Grammatical forms represent categories and relationships (for example, nouns name categories of natural kinds while verbs may depict intracategory relationships).
7. For reading to occur, categories of information in text must map onto categories of information in mind. Words are the medium through which this transaction occurs.
8. Generally, feature analysis is used to generated taxonomies when thematic relationships are not used.

Limitations

1. The effect of other variables was not controlled. These might include cognitive variables such as I.Q. or cultural variables such as class background.
2. The results of this investigation cannot be generalized to populations other than that which was used in the research.
3. The results of this investigation cannot be generalized to standardized reading tests other than the one which was used in the research.

Delimitation

This study is delimited to students enrolled in the 1991/91 school year in grades one and three in a rural east Tennessee county.

APPENDIX D

RAW DATA

Table IX

Raw Data - First Grade
Low Reading Group
By Subject

Sub- ject	Sta- nine	Defining Features	Proto- types	Natural Kinds	Super- ordinates	Taxon- omies	Themes
1	2	0	2	1	0	11	0
2	2	0	1	3	0	4	0
3	3	0	3	3	0	3	0
4	2	0	1	2	0	1	0
5	3	0	4	0	0	5	1
6	2	0	0	3	0	7	0
7	2	0	3	4	0	0	0

Table X

Raw Data - First Grade
High Reading Group
By Subject

Sub- ject	Sta- nine	Defining Features	Proto- types	Natural Kinds	Super- ordinates	Taxon- omies	Themes
1	9	1	1	3	1	6	0
2	8	0	3	1	0	3	1
3	8	1	1	4	0	3	0
4	7	1	3	2	0	17	1
5	7	1	2	2	0	3	0
6	8	1	3	4	0	7	0
7	7	0	2	3	0	10	0
8	7	0	4	4	0	3	0
9	7	0	3	2	0	5	2

Table XI

Raw Data - Third Grade
Low Reading Group
By Subject

Sub- ject	Sta- nine	Defining Features	Proto- types	Natural Kinds	Super- ordinates	Taxon- omies	Themes
1	3	0	1	2	0	4	1
2	2	0	1	5	0	9	0
3	3	0	0	3	0	11	0
4	3	0	2	3	0	9	0
5	3	0	1	4	0	4	0
6	1	1	2	4	1	15	0
7	1	0	0	4	1	0	0
8	1	0	4	2	0	3	0
9	3	1	3	2	0	15	0
10	3	0	2	2	0	12	0

Table XII

Raw Data - Third Grade
High Reading Group
By Subject

Sub- ject	Sta- nine	Defining Features	Proto- types	Natural Kinds	Super- ordinates	Taxon- omies	Themes
1	8	1	3	4	1	6	0
2	8	0	4	5	0	5	2
3	7	1	2	4	1	6	0
4	8	0	1	4	0	11	0
5	7	1	3	3	1	20	0
6	7	1	2	3	0	22	0
7	8	0	3	4	0	7	0

VITA

Robert Rogers Wheeler II was born in Ann Arbor, Michigan on March 14, 1947. He attended elementary and high school in Grand Rapids, Michigan. Following graduation, Mr. Wheeler enrolled in Grand Valley State University. He received his Bachelor of Science degree in sociology and psychology with elementary education certification in the spring of 1969.

Mr. Wheeler taught the sixth grade in Grand Rapids, Michigan prior to enrolling in a graduate program. He received his Master of Arts degree in sociology of education with a collateral in psychology in the spring of 1974. After graduation, Mr. Wheeler taught the third grade in Grand Rapids, Michigan.

Mr. Wheeler's parallel careers in business and education required several relocations in Michigan and Florida. In 1985 he moved his family to Tennessee and enrolled in the University of Tennessee one year later. Mr. Wheeler attended the University and was a graduate assistant from 1988 to 1990. During this time, he served as a clinician and teacher. In May of 1992 he received the degree of Doctor of Education in Curriculum and Instruction. His major work was in reading and language arts, with collaterals in psychology and research.

He is a member of Phi Delta Kappa, The International Reading Association, and the Tennessee Reading Association.

Mr. Wheeler lives in Gatlinburg, Tennessee with his wife, Linda, and their two children, Sugar and Taffy.